开創新世紀的
数學文化

陳省身

二千年十一月

译丛序言

数学,这门古老而又常新的科学,已阔步迈进了21世纪.

回顾过去的一个世纪,数学科学的巨大发展,比以往任何时代都更牢固地确立了它作为整个科学技术的基础的地位.数学正突破传统的应用范围向几乎所有的人类知识领域渗透,并越来越直接地为人类物质生产与日常生活作出贡献.同时,数学作为一种文化,已成为人类文明进步的标志.因此,对于当今社会每一个有文化的人士而言,不论他从事何种职业,都需要学习数学,了解数学和运用数学.现代社会对数学的这种需要,在未来的世纪中无疑将更加与日俱增.

另一方面,20世纪数学思想的深刻变革,已将这门科学的核心部分引向高度抽象化的道路.面对各种深奥的数学理论和复杂的数学方法,门外汉往往只好望而却步.这样,提高数学的可接受度,就成为一种当务之急.

一般说来,一个国家数学普及的程度与该国数学发展的水平相应并且是数学水平提高的基础.随着中国现代数学研究与教育的长足进步,数学普及工作在我国也受到重视.早在60年代,华罗庚、吴文俊等一批数学家亲自动手撰写的数学通俗读物,激发了一代青少年学习数学的兴趣,影响绵延至今.改革开放以来,我国数学界对传播现代数学又作出了新的努力.但总体来说,我国的数学普及工作与发达国家相比尚有差距.我国数学要率先赶超世界先进水平,数学普及与传播方面的赶超乃是一

个重要的环节和迫切的任务.为此,借鉴外国的先进经验是必不可少的.

《通俗数学名著译丛》的编辑出版,正是要通过翻译、引进国外优秀数学科普读物,推动国内的数学普及与传播工作,为我国数学赶超世界先进水平的宏伟工程贡献力量.丛书的选题计划,是出版社与编委会在对国外数学科普读物广泛调研的基础上讨论确定的.所选著述,基本上都是在国外已广为流传、受到公众好评的佳作.它们在内容上包括了不同的种类,有的深入浅出介绍当代数学的重大成就与应用;有的循循善诱启迪数学思维与发现技巧;有的富于哲理阐述数学与自然或其他科学的联系;等等,试图为人们提供全新的观察视角,以窥探现代数学的发展概貌,领略数学文化的丰富多采.

丛书的读者对象,力求定位于尽可能广泛的范围.为此丛书中适当纳入了不同层次的作品,以使包括大、中学生;大、中学教师;研究生;一般科技工作者等在内的广大读者都能开卷受益.即使是对于专业数学工作者,本丛书的部分作品也是值得一读的.现代数学是一株分支众多的大树,一个数学家对于他所研究的专业以外的领域,也往往深有隔行如隔山之感,也需要涉猎其他分支的进展,了解数学不同分支的联系.

需要指出的是,由于种种原因,近年来国内科技译著尤其是科普译著的出版不景气.在这样的情况下,上海教育出版社按照国际版权公约,不惜耗资购买版权、组织翻译出版这套《通俗数学名著译丛》,这无疑是值得称道和支持的举措.参加本丛书翻译的专家学者们,自愿抽出宝贵的时间来进行这类通常不被算作成果但却能帮助公众了解和欣赏数学成果的有益工作,同样也是值得肯定与提倡的.

像这样集中地翻译、引进数学科普读物,在国内还不多见.值得高兴的是,这项工作从一开始就得到了数学界许多人士的赞同与支持,特别是数学大师陈省身先生两次为丛书题词,使我

们深受鼓舞.到目前为止,这套丛书已出版了13种,印数大多逾万,有的已经是第四次印刷,这对编译者来说确是令人欣慰的信息.我们热切希望广大读者继续关心、扶植这项工作,使《通俗数学名著译丛》的出版获得更大的成功.

让我们举手迎接数学科学的新的黄金时代,让公众了解、喜爱数学,让数学走进千家万户!

《通俗数学名著译丛》编委会

2001年8月

序　言

这本书主要由《数学信使》(*Mathematical Intelligencer*)上"数学娱乐"(Mathematical Entertainments)栏目中的文章组成,我于1991年至1996年主持编辑了这个特色专栏. 这个栏目本来几乎是专门用于讨论题目的,但当我接手时,在这本杂志主编阿克斯勒(Sheldon Axler)的鼓励下,我开始按我喜欢的方式来进行编辑. 结果是题目变得越来越少,最后全部消失. 相反,我却能撰写任何令我感兴趣的东西. 结果是选题多种多样,遍及这块领地,不过几乎一直都在关注着非常初等的东西. 例如,有三章差不多专门是讨论三角形的,还有两章是讨论用矩形进行铺砌的,有三章多是讨论由简单递归关系给出的序列的神秘性质的,三章是讨论游戏和悖论的,三章是讨论一种特定的自动机的.

从这些列举的这些内容可以看出,似乎并没有把所收集的材料统一起来的主线,但事实上我在决定选用哪些材料时,遵循着某种一般的原则.

(1) 粗略地说,不应当要有任何专门化的知识才能理解所选的论题(虽然偶尔有一些随后的分析会变得稍稍超前一些.)

(2) 高度优先地考虑数学上的意外发现和没有料到的转折,其中有些得到了解释,但还有一些仍然莫名其妙.

这两条原则都通过来自日常生活但竟然具有数学内容的例子得到了体现. 其中有我们熟悉的儿童游戏,还有两个章节是关于系鞋带的. 在同样的精神下,在一个关于几何作图的章节

中，人们居然可以扔掉使用不灵便的圆规，代之以一把可以在上面标上记号的直尺．这个工具让人们不仅可以完成传统的作图，而且可以完成任何涉及最高为4次的方程的作图，从而使得三等分角和倍立方成为可能(见第10章)．在前三个附录中，我提供了一些材料的最初来源．

计算机在许多章节中都有表演，但扮演的角色各种各样．有时候它们只是用来执行计算，例如找出2的几次方幂其开头4位数字是1992(第7章)．在其他地方它们用来进行实验，而最为令人感兴趣的或许是证明定理．例如，人们通过数值计算发现了某些在几何上定义的点是共线的，从而得到了几十个平面几何的新定理(第6章)．当这些被发现的共线性接下来用代数式子表示出来时，计算机又以它的符号运算模式而不是数值计算模式参与其事，提供了证明．

然而，以这本书的观点，使用计算机所导致的最令人着迷的后果是，它对我们思考数学本身的方式产生了影响．从传统上讲，数学是一种演绎科学的**典型**例子．例如，不管人们对——比方说——费马大定理的正确性有多少“实验证据”，但在这个定理被证明之前，也就是说，在被证明能由先前所知道的结果推导出来之前，它并不被认为是数学．但是在1931年，事情开始发生变化．那一年哥德尔证明了不存在非平凡的数学系统能够证明或否证所有可在这系统中作出的陈述．今天，由于让我们能聚集海量的经验数据，计算机使这场危机变得十分明朗．例如，根据最新的计算，我们现在知道了π的几十亿位小数，其中大约有十分之一是7，而且偏差非常之小．然而怎样证明这个行为将“永远”继续下去，没有人能拿得出主意．事实上，说不定从某一位开始根本就没有7了．乐观主义者会说一定有一个证明在其他什么地方，而当我们变得足够聪明的时候，我们就会发现它．但要是根本就不存在证明呢？在第3章和第5章中，我们考察了另一类关于序列的问题，这里的某些结果看起来极有可能成

立，然而找到证明的希望又是极其的渺茫．于是，数学家们发觉自己处于一种奇特的境地．在一个极端，通过聚集数据，计算机能够先是为我们提供猜想，然后用我们设计的程序，进一步解决它们．上面提到的几何结果就是这种情况．在另一个极端，计算机可能是在让我们看一下正确的数学事实，但对这些事实根本不存在什么证明．当然，在这两个极端之间，还有着足够的空间让数学家为那些可预见的甚至不可预见的未来忙个不停．

因此计算机的到来使数学更像物理科学，因为现在它除了有理论成分外，还有强大的实验成分．为发现物理世界中所发生的情况，物理学家使用粒子加速器，而天文学家使用高能望远镜．对数学家来说，计算机现在扮演了一个类似的角色．重要的区别在于，计算机给我们的信息不是关于物理世界的，而是关于"抽象"世界的．我们研究的东西，有限群，解析函数，其他等等，它们的"真实"性并不比原子或星系更差．区别在于这些对象是抽象的概念，是我们不把它们发明出来就不存在的东西．当然，这些抽象概念中有许多最初产生于一种试图理解物理世界的努力之中，一个典型的例子就是欧几里得的著作．但是也存在着所谓的"纯粹"数学，其中研究的对象就是这些抽象概念本身，因此这里有着一个很大的悖论．作为一个实例，让我们再考察一下 π 的小数，看一看在描述这个问题时涉及了什么．首先我们需要一个圆的概念．大自然给出了提示性的例子：太阳、月球、蘑菇头．但是大自然没有提供圆心．我们得把它发明出来，同样的还有圆周和直径，以及作为一个"数"（这也是我们的发明）的长度概念，最后还有十进制展开．所有这些东西都是人类心智的创造，然而它们似乎具有它们自己的生命，因为它们站起身来向我们提出谜题，其中有些可能我们永远也不能解决．本书的最后一个附录进一步探索了这种事态的哲学涵义．

最后，作一个澄清．正如所说的，这个栏目最初是作为"娱乐"而设立的，而我上面也提到了向我们提供谜题的数学现象．

这些词的通常涵义让人联想到我们在这里只是涉及一些轻松的事情，而不是那种“严肃的数学”．这种区分只是表面上的而不是实际上的．事实上，人们可以说整个科学事业都是在解决大自然不断抛给我们的谜题．然而，我试图只表现出其中轻松的一面是从某种意义上来说的．努力理解数学结果的过程可能是一种非常艰苦的劳动．对比之下，在这本书中，我努力选取了这样的一些论题：读者用至多中等程度的精力消耗就能从中获得作为回报的发现，换句话说，达到最大的快乐努力比．虽然这本书不打算做到读起来像小说那样，但我的希望是，通过在书中深入挖掘，读者将发现令人愉悦的东西，而且，也是富有启发性的东西．

戴维·盖尔
于加利福尼亚大学
伯克利校区

致 谢

我要向那些为本书作出贡献的人们表示感谢，他们在帮助把这些材料组织起来编成本书的过程中起了重要的作用.

我要感谢《数学信使》的两位主编阿克斯勒和戴维斯(Chandler Davis)对我的支持和鼓励，他们在我担任“数学娱乐”编辑期间，让我放手撰写任何令我感兴趣的东西. 有些章节的部分或全部是由“客座”专栏作家撰写的，其中包括纽曼(Donald Newman)、普罗普(Jim Propp)、萨瑟兰(Scott Sutherland)、特罗别茨科伊(Serge Troubetzkoy)、斯坦(Sherman Stein)、霍尔顿(John Halton)、戈隆布(Solomon Golomb)、米休列维奇(Michal Misiurewicz)和马查多(Armando Machado).

此外，许多章节的材料是通过与许多人的讨论、电子邮件通信和电话交谈而获得的，他们中有索莫斯(Michael Somos)、巴拉尼(Imre Barany)、金伯林(Clark Kimberling)、伯利坎普(Elwyn Berlekamp)、凯勒(Joe Keller[①])、克利福德·加德纳(Clifford Gardner)，特别是已故的鲁滨逊(Raphael Robinson)，他对几乎所有的早期章节都作出了关键性的贡献.

所有这些人使得那些专栏文章的写作和最后的成书成为一种合作风险事业. 我希望读者将分享到一些我们在编这本书时所感到的愉悦.

---

① 本书第 17 章提到一位 K. Joy，疑为此人. ——译注

目 录

第1章 令人困惑的简单序列

计算机产生的谜团

许多数学家认为,计算机对数学的主要影响,并不是像有人可能料想的那样在于解答问题,恰恰相反,而是在于提出问题.最近在离散动力系统方面的工作可能是最好的说明,这一领域的研究是由费根鲍姆(Mitchell Feigenbaum)那著名的计算机实验[①]所激发的.或许"探测"是对这些工作的一个较好的描述,因为要对它们作出恰当的比喻,不能用物理学或者生物学,而要用天文学.计算机是数学家的天文望远镜,对它的巧妙运用使得人们能够发现在数学宇宙中那"遥远的地方"有些什么.(这一发展从整体上应该让柏拉图(Plato)主义者觉得满意,他们一直在说,数学现象就如同恒星和星系那样是被发现的而不是被发明的.)

这里所描述的最新工作又提供了一个例子,因为它们揭示了一系列令人意外的现象,这些现象如果不使用计算机恐怕是永远也不会被观察到的. [1]

---

① 这是指美国洛斯阿拉莫斯国家实验室的物理学家费根鲍姆于1975年通过计算机发现了关于混沌行为的普适常数 $\alpha = 4.6692\cdots\cdots$ 一事. 关于这个普适常数的意义,可参见本译丛中《数学:新的黄金时代》的第4章和《计算出人意料》的附录2. ——译注

索莫斯序列传奇

索莫斯(Michael Somos)在研究椭圆 θ 函数的性质时发现了一个无穷序列,它的开头 15 项为

1,1,1,1,1,1,3,5,9,23,75,421,1103,5047,41783.

这序列由 $a_i = 1(0 \leqslant i \leqslant 5)$ 和

$$a_n = (a_{n-1}a_{n-5} + a_{n-2}a_{n-4} + a_{n-3}^2)/a_{n-6} \quad (n > 5) \quad (1)$$

所定义. 令人惊奇的是,就我们的眼睛 = 计算机 = 天文望远镜所能看到的,这个递归关系总是产生整数. 事实上,对这个例子来说,并不需要一架天文望远镜. 一副好的双筒望远镜就够了. 例如,用一个袖珍计算器你就能很容易地验证下一项的分子能被 23 整除,因此 a_{15} 仍然是一个整数. 接下去的情况又是如何呢[①]?

在对这一现象进行观察的基础上,许多人想到应考察较为简单的 4 阶递归关系

$$a_n a_{n-4} = a_{n-1}a_{n-3} + a_{n-2}^2, \ a_0 = a_1 = a_2 = a_3 = 1. \quad (2)$$

结果所有的项还是都为整数,但是在这种情况下局面是可以控制的. 一些人已经得到了证明,第一个证明是马洛夫(Janice Malouf)给出的. 我们在这里介绍一个不同的证明,它是由伯格曼(George Bergman)给出的. 首先注意到这个序列的任何连续 4 项都两两互素. 因为假设这个结论对 a_n 以前的各项都成立,那么当且仅当 a_n 与 a_{n-1} 或 a_{n-3} 有一个公共素因数 p 时,p 也整除 a_{n-2},这就同归纳法假设发生了矛盾.

现在我们用归纳法证明如果 $a_{n-4},\cdots,a_n,\cdots,a_{n+3}$ 都是整数(显然这对 $n = 4$ 成立),则 a_{n+4} 也是整数,从而所有的 a_i 都是整数. 记 $a_{n-3} = a, a_{n-2} = b, a_{n-1} = c$,我们有 $a_n a_{n-4} = ac +$

---

① 其实早在 14 年前索莫斯就发现了这个序列,但直到 1989 年夏才引起数学界的注意. —— 原注

b^2，因此 a_n 整除 $ac+b^2$. 根据上一段所得的结论，我们用(2)对这个序列进行模 a_n 的同余式运算[①]后可以得出

$$a,b,c,0,\frac{c^2}{a},\frac{c^3}{ab},\frac{c^3}{a^2},a_na_{n+4}\equiv\frac{c^5}{a^3b^2}(ac+b^2)\equiv 0. \quad (3)$$

因此 a_n 整除 a_na_{n+4}.

虽然这个证明非常简单，但是它依赖于第4次迭代出现了因式 $ac+b^2$ 这个碰巧发生的事实. 我们在后面将再提到这一点.

这个方法同样适用于5阶递归关系

$$a_na_{n-5}=a_{n-1}a_{n-4}+a_{n-2}a_{n-3}. \quad (4)$$

事实上，在所有这些递归关系中，人们可以用任何整数来作为 $a_{n-i}a_{n-j}$ 项的系数，结果仍然得出整数，而且能够证明这对递归关系(2)和(4)成立[②].

[2]

下一步进展来自希克森(Dean Hickerson)，他证明了原来那个索莫斯序列总是给出整数. 事实上，他证明了更广的一些事情. 他考虑这个序列时不是以6个1为初始值，而是以不定元 $a_0,a_1,\cdots,a_5$ 为初始值. 这样，这个递归关系就产生了这些 a_i 的有理函数 $a_n=p_n/q_n$，而他的定理是说，这些函数的分母总是系数为1的单项式. 这对 $a_0,\cdots,a_{11}$ 是很清楚的，但请注意，要计算 a_{12}，人们必须除以 $a_6=(a_5a_1+a_4a_2+a_3^2)/a_0=p_6/a_0$. 对 a_7,a_8 等项的手工计算很快就变得无法驾驭. 但这正是符号运算程序的用武之地. 希克森用 *Macsyma* 软件发现，就像上面

---

① 关于同余式运算，可参见本译丛中《数论妙趣》的第5章或其他初等数论教材. 这里要注意的是，由于 a_n 与 a,b 分别互素，所以存在整数 a',b'，使得 $aa'\equiv 1(\mathrm{mod}\ a_n)$ 和 $bb'\equiv 1(\mathrm{mod}\ a_n)$. 一般记 $a'=a^{-1},b'=b^{-1}$. 但这里用了分式来表示，因此下面的序列中，$\frac{c^2}{a}$，$\frac{c^3}{ab}$，$\frac{c^3}{a^2}$ 和 $\frac{c^5}{a^3b^2}$ 实际上代表着整数. —— 译注

② 如果允许这些系数为负整数，则可能发生某个 $a_n=0$ 的情况，这时我们按惯例认为这个序列至此终止. —— 原注

的(3)那样，p_6 作为一个因式出现在 a_{12} 的分子(把它化简后还包括194项!)中. 他用 *Macsyma* 进行计算还证明了 p_6 与 p_7，…，p_{12} 互素，并用归纳法完成了证明.（斯坦利(Richard Stanley)用类似的方法也解决了这个问题.）

但是我们领会到了什么?正如希克森所说的，“关于我的证明，我所不喜欢的是，它没有解释为什么这个结果会成立. 它主要依赖于这样一个事实：当你计算 a_{12} 时，发生了出乎意料的消项. 但是为什么会发生这种情况呢?”确实，这个证明与其说阐明了这种现象，还不如说(如果真有什么作用的话)使这种现象更为神秘. 我报告这件事时带着某种困窘，因为我早先断言数学中的一个证明在某种意义上等价于一种解释. 我们现在看到显然不一定是这么一回事. 或许，当我们找到(如果能够找到的话)一个“正确的”证明时，情况会变得清晰起来，但是必定会存在一个正确的证明吗? 这使我们想起四色定理的证明[①]. 索莫斯问题的最令人感兴趣的特点之一，就是它把我们引向这类思索.

回到手中的问题. 人们已经发现了关于4阶，5阶，6阶递归关系的证明和关于7阶递归关系的经验证据，但是我们发现8阶和8阶以上的递归关系并不总是给出整数. 你可以用袖珍计算器很容易地验证这一点. 例如，在8阶递归关系中，a_{17} 是一个分数. 真是越来越古奇[②].

---

① 关于四色定理及其计算机证明，可参见本译丛中《数学：新的黄金时代》的第7章.—— 译注

② 原文为 Curiouser and curiouser，出自19世纪的世界著名儿童文学《爱丽丝漫游奇境记》(*Alice's Adventure in Wonderland*). 这是一个不合英语规范的句子，是书中主人公爱丽丝因遇到一连串怪事而惊奇得连话都说不好时发出的惊叹. 正确的说法应是 More and more curiouse. 这里也用不规范的中文译出. 该书作者卡罗尔(Lewis Carroll)，真名道奇森(Charles Lutwidge Dodgson)，英国牛津大学数学讲师. 现代数学普及作家常引用该书中的情节和句子. —— 译注

接下来的发现是由鲁滨逊(Raphael Robinson)得到的,他发现递归关系(1),(2)和(4)的这种整数性(看起来)也为一个无穷的递归关系族所拥有.即对任何的 $k \geqslant 6$,从 k 个1开始,然后利用递归关系

$$a_n a_{n-k} = a_{n-1} a_{n-k+1} + a_{n-2} a_{n-k+2} \tag{5}$$

或

$$a_n a_{n-k} = a_{n-1} a_{n-k+1} + a_{n-2} a_{n-k+2} + a_{n-3} a_{n-k+3}. \tag{5'}$$

现在人们要对付的是一个有无穷多个序列的集合,这似乎把问题推出了 *Macsyma* 式证明力所能及的范围.

在这个时候,我的袖珍计算器使我确信,对任何的 $0 < l < m < k$,递归关系

$$a_n a_{n-k} = x a_{n-l} a_{n-k+l} + y a_{n-m} a_{n-k+m} \tag{6}$$

总是给出整数,这就推广了(5).仍然是鲁滨逊的进一步研究,导致了下面的

猜想 对于任何的 $p, q, r < k$,当且仅当 $p + q + r = k$ 时,递归关系

$$a_n a_{n-k} = x a_{n-p} a_{n-k+p} + y a_{n-q} a_{n-k+q} + z a_{n-r} a_{n-k+r} \tag{7}$$

总是给出整数.

[3]

(鲁滨逊的证据仅适于 $x = y = z = 1$ 时的情况.加上任意的 x, y, z 则是我的主意.)这样就把(5′)和(6)归并了进来,具体地说,(6)对应于 $p = l, q = k - m, r = m - l$ 以及 $z = 0$ 的情况,而(5′)对应于 $p = 1, q = 2, r = k - 3$ 以及 $x = y = z = 1$ 的情况.

故事还没有完.斯科特(Dana Scott)对 $k = 4$ 这种最简单的情况设计了一个程序,但忘了对 a_{n-2} 项平方.然而这递归关系仍给出整数!事实上,人们发现递归关系(2)可以推广为

$$a_n a_{n-4} = a_{n-1}^p a_{n-3}^q + a_{n-2}^r, \quad 对任何\ p, q, r > 0. \tag{8}$$

而且伯格曼的证明对任意的指数情况都有效,就像它在对递归

关系(4) 时那样. 另一方面,人们不能**同时**对系数和指数进行随意的选择. 事实上,递归关系 $a_n a_{n-4} = 2a_{n-1}a_{n-3} + a_{n-2}$ 并不总是给出整数(虽然如果左边是 $a_n a_{n-3} + y a_{n-2}$,则可以证明对任何的 y 这递归关系总是给出整数).

递归关系(8) 十分有趣,因为在其他所有例子中左边都是齐次的. 这是不是节外生枝?或许是的,但当我们处理递归式左边有 3 项的序列时,就再也不能随意取指数. 事实上,如果我们在最初的序列(1) 中忘了对 a_{n-3} 平方,就会得出分数.

或许所有这类递归关系中最简单的已为斯科特所发现. 即对任何的 k,

$$a_n a_{n-k} = a_{n-1}^2 + \cdots + a_{n-k+1}^2. \tag{9}$$

它看来对所有的 k 都产生整数. 其他的"好" 递归关系看来有

$$a_n a_{n-k} = a_n a_{n-2} + \cdots + a_{n-k+2}a_{n-k+1}\text{①}, \tag{10}$$

以及,对 k 为奇数,

$$a_n a_{n-k} = a_{n-1}a_{n-2} + a_{n-3}a_{n-4} + \cdots + a_{n-k+2}a_{n-k+1}. \tag{11}$$

这些递归关系开辟了一片新天地,因为它们左边的项数可以任意多,而在前面的例子中看来最多只有3项. 对 $k = 4$,伯格曼的证明对(9) 和(11) 有效,但是对(10) 无效. 后者(暂时) 还没有得到解决.

我不想把这个话题无限制地扯开去,因为看来这些索莫斯序列的新例子出现的速度比我能把它们写下来的速度要快. 有一个完整的领域,其中人们使用了如同(1)那样的递归关系,但初始值序列不是全为 1,而是既有 1 又有 2,或既有 1 又有 -1,等等. 实验显示有时人们得到整数而有时并不,看来不存在可以辨认的模式. 在积极的方面,利用希克森的思想,盖尔(David

---

① 原文如此,疑有误,因为右边也出现了 a_n,而且从右边仅列的两项也看不出省略号所代表的其他各项应该是什么样. 估计可能是 $a_n a_{n-k} = a_{n-1}a_{n-2} + a_{n-2}a_{n-3} + \cdots + a_{n-k+2}a_{n-k+1}$. ——译注

Gale)和鲁滨逊证明了序列(5)(但不是(5′))的整数性. 我强烈地预感到,到我这篇文章印出发表的时候,关于这些索莫斯序列人们将会知道得更多. 或许问题甚至已被解决,但在写这篇文章的时候,情况还是扑朔迷离.

补遗 这篇专栏文章写完后没有几个星期,洛托(Ben Lotto)利用希克森的思想,但没有用计算机,证明了(9)总是给出整数. 然而,他的方法看来对(10)和(11)无效,虽然鲁滨逊用一种完全不同但初等的方法,证明了(10)在 $k = 4$ 时总给出整数,从而解决了上面再前一段文字中提出的问题. 还有,猜想(7)在 $k = 7, p = q = r = 1$ 时的情况已得到证明①(这次用的是 *Mathematica* 而不是 *Macsyma*). 最后,鲁滨逊发现当把这些序列中各项的值以模 n 取剩余后②将出现一系列周期性现象. 这种周期性对(2),(4),以及 $k = 4$ 时的(10)和 $k = 3$ 时的(9)已得到证明,但对其他情况还不能解释. [4]

伦斯特拉③用数学定理开的玩笑

"不存在完满平方④. 假设 n 是一个完满平方. 考察 n 的奇因数. 它们都整除它们中最大的数,而这个最大的奇因数本身

① 原文如此, 疑有误,估计应是 $k = 7, x = y = z = 1$ 时的情况.—— 译注

② 即用 n 除后取余数. —— 译注

③ 伦斯特拉(Hendrik Lenstra),当代美国数学家,以用计算机研究数论问题而闻名.——译注

④ 原文为 perfect square. 一般译为"完全平方",这里按全国自然科学名词审定委员会1993年公布的《数学名词》,译为"完满平方",指能表示为某一整数平方的数,如 $4 = 2^2$,或指能表示为某一多项式平方的多项式,如 $x^2 + 4x + 4 = (x + 2)^2$. 但下文的证明却把这个词理解为既是完满数(又译"完全数")又是完满平方的数. 所谓完满数,是指其所有因数(包括其本身)之和为此数两倍的数,如6,其所有因数之和 $1 + 2 + 3 + 6 = 12$(可参见本译从中《数论妙趣》的第3章). 附带说一下,perfect square 还指"完美正方形",即由大小不同的若干个小正方形拼成的一个大正方形(可参见本书的第9章和第12章). ——译注

是一个平方数,记为 d^2. 这说明 n 的奇因数总是成对出现,即奇因数 a 与 b 以 $a \cdot b = d^2$ 的关系配对. 只有 d 同它本身配对. 因此 n 的奇因数的个数为奇数. 这意味着 n 的所有因数之和为奇数. 特别是,不会等于 $2n$. 因此 n 不是完满数. 这就产生了一个矛盾.所以不存在完满平方."

明白了吗?

评注 看来这个玩笑只适于用英语. 在其他语言中,平方就是平方(然而这定理却是国际通用的).

是否有一种数学基因

一位数理逻辑学家的 4 岁外甥女正在玩一个游戏. 她扮演火车上的列车员,她的母亲扮演一位乘客. "等一下,"南茜说,"我们必须弄一些纸片来做车票." "哦,"母亲说,她可能已经劳累了一天,"我们非得这样吗? 毕竟这只是个用假车票玩的假游戏." "哦不,妈咪,你错了,"南茜答道,"车票是假的,但游戏却是真的." [5]

[6]

第2章 概率论悖论

悖论与一对盒子

什么是悖论？或许数学中最著名的例子是罗素悖论[①]和巴拿赫－塔尔斯基悖论[②]，但是这两个结果十分不同．巴拿赫－塔尔斯基定理被认为是悖论，是因为它表明集合的行为可以与我们对它们的直觉观念有很大的区别．而在另一方面，罗素悖论则表明，从看起来是合理的公理出发，人们可能会导出一个矛

---

① 这个由著名英国哲学家罗素(Bertrand A. W. Russell)于1901年提出的悖论内容如下：如果一个集合 x 不以它本身作为元素，即 $x \notin x$，则称 x 为正常集合，否则称 x 为异常集合．现令 X 为由一切正常集合所组成的集合．若 $X \in X$，则由 X 的组成原则，X 是正常集合，但由正常集合的定义，$X \notin X$；若 $X \notin X$，则由 X 的组成原则，X 是异常集合，但由异常集合的定义，$X \in X$．无论哪种情况，均导致矛盾．罗素悖论触发了数学的第三次危机．后数学家们(包括罗素本人)提出了各种集合论公理系统，避免了这个悖论．有兴趣的读者可参见《第三次数学危机》(胡作玄著，四川人民出版社，1985年版)．——译注

② 这个由著名波兰数学家巴拿赫(Stefan Banach)和塔尔斯基(Alfred Tarski)于1924年提出的悖论又称巴拿赫－塔尔斯基分球定理．它在直观上等于说：任意一个球体都可以被分成有穷多份后重新组合成两个与原来一样大小的球体．如此怪异的结论竟然在数学上可予以严格的证明．其中的症结在于人们接受了一个在直觉上十分自然的公理，即所谓“选择公理”：对于一族两两不相交的非空集合，总可以从每个集合中各选出一个元素组成一个新的集合．详情可参见《超穷数与超穷论法》(谢邦杰编著，吉林人民出版社，1979年版)，但要求读者具有一定的现代数学知识．——译注

盾.它的恰当名称不是悖论(paradox),而是**二律背反**(antinomy).根据韦氏大词典,二律背反是指"在两个显然同样令人信服的原理之间的,或者在从这些原理正确地导出的推论之间的矛盾",而悖论则是"看起来与常识矛盾或者对立的,然而可能是正确的命题".

下面前两个例子是二律背反,而后两个是悖论.

1.另一个盒子 在你面前有两个封闭的盒子,每个盒子里都有一定数量的钱.这些钱当初是按以下规则放进去的.连续抛掷一枚质量均匀的硬币,直到它落下来反面向上为止.如果连续掷了 n 次落下来都是正面向上,到第 $n+1$ 次才反面向上,则在一个盒子里放 3^n 美元,而在另一个盒子里放 3^{n+1} 美元.现在允许你打开其中一个盒子,数一数里面有多少钱.你可以把这些钱放进自己的口袋,也可以改变主意,拿走另一个尚未打开的盒子里的钱.你该怎么办?很显然,如果你打开的盒子里面只有1美元,那么你应该拿走另一个盒子里的3美元.现在假定你打开的盒子里面有 3^n 美元,那么容易看出,另一个盒子里将分别以2/3和1/3的概率有着 3^{n-1} 美元和 3^{n+1} 美元.于是你换一个盒子所能得到的钱的数学期望为[7]

$$\frac{2}{3}3^{n-1}+\frac{1}{3}3^{n+1}=\frac{11}{9}3^n>3^n,$$

因此改取另一个盒子里的钱,就可以使你所得钱的数学期望达到最大.假定你是个按最大数学期望行事的人,这就是你将采取的行动.(在按最大数学期望行事是否"合理"的问题上存在着争议,但这显然是另外一回事,因为我们在这里关心的仅是数学而不是它对人的行为的意义.)但是现在既然你已事先知道你总要改取另一个盒子里的钱,那么就没有理由浪费时间去打开一个盒子并数一数里面的钱了.你应该一开始就直接选"另一个盒子".然而用同样的道理可以证明,无论你选哪一个盒子,从数学期望的角度看,你还是选另一个盒子为好.

这个游戏有着某种彼得堡悖论[①]的风格．但在那个悖论中，所得钱的数学期望为无穷大．而我们这个不同版本的新颖之处在于，它看来导出了一个矛盾，因此我们现在面临的是一个二律背反而不是一个悖论．

2．打败赌场老板 杜宾斯(Lester Dubins)给我讲了一个有点类似的例子，但他不清楚它的来源．在某个娱乐性赌场，人们可以玩这样一种赌博游戏．赌场老板公布一个正整数 n．在这个赌博中，掷一枚质量均匀的硬币直到它反面向上为止这件事是请你这位赌客来做的．如果你掷了 $n-1$ 次，你就输给老板 8^{n-1} 美元，但如果你掷了 $n+1$ 次，则你就从老板那儿赢得 8^n 美元．所有其他情况都算平手．因为恰好掷了 n 次的概率是 $\frac{1}{2^n}$，你所赢钱的数学期望就是 $\frac{8^n}{2^{n+1}}-\frac{8^{n-1}}{2^{n-1}}=4^{n-1}(n>1)$ 或 $2(n=1)$，因此你所赢钱的数学期望，也就是老板所输钱的数学期望，是一个正数．但是现在发现赌场老板原来是这样来确定 n 的：他也是掷同一枚质量均匀的硬币，直到第一次出现反面向上为止，如此所掷的次数就确定为 n．这样，你和老板就是以完全对称的方式来玩这个赌博的．你们每人都按以上方式掷硬币，如果两人所掷的次数恰好是两个相邻的整数 n 和 $n+1$，那么掷了 n 次的那位就付 8^n 美元给掷了 $n+1$ 次的那位．但是我们刚才看到，不管老板宣布的是哪个数，从数学期望的角度测算，这个赌博是

---

① 这是在概率论发展史上十分著名的一个悖论，其内容如下：设有一种赌博，赌客需先付一笔"入场费"方可参加．赌法是：赌客以文中所述方式掷硬币，如果他掷了 r 次后第一次出现反面向上，则他从赌场老板那儿赢得 2^r 元．现在问：要使这场赌博显得"公平"，赌客应付多少入场费？显然，这入场费应等于赌客所赢钱的数学期望．但计算一下即可知，这个数学期望为 $\frac{1}{2}\cdot 2+\frac{1}{2^2}\cdot 2^2+\cdots+\frac{1}{2^r}\cdot 2^r+\cdots=1+1+\cdots+1+\cdots=\infty$．关于这个悖论的详细讨论，可参见《概率论及其应用(下册)》(W．费勒著，刘文译，科学出版社，1979年版)．——译注

有利于你的．在一个完全对称的赌博中怎么会出现这种不对称性呢？

3．还是另一个盒子（对布莱克韦尔（David Blackwell）所提出的一个例子稍作修改而得）　这次盒子里放的不是钱，而是每个盒子里放一个整数（可能是印在一张卡片上）．而你所知道的唯一事情是它们不相同．你从中随机地挑选一个，然后你得猜一下另一个数比你所选的那个数是大还是小．你能做一些什么事使得你猜得比猜对猜错各占一半机会的情况更好一些吗？出人意料的是，答案是能，条件是你要有某种产生随机性的工具，例如用来抛掷的一枚质量均匀的硬币．假定你有一个指针式转盘，就像在儿童玩的棋盘游戏中所用的那种．你应该先转一下转盘，然后记下从指针初始位置到终止位置之间的夹角 θ．现在

[8] 你挑选一个数，然后按 $\cot\frac{\theta}{2}$ 比你所选的数是大还是小来猜另一个数是大还是小．为方便起见，让我们假设 θ 在 $(0,2\pi]$ 上一致分布．我们断言：如果那两个数是 $p > q$，那么你猜对的概率就是 $\frac{1}{2}+\frac{\operatorname{arccot}q-\operatorname{arccot}p}{2\pi}$．具体证明如下：当且仅当 $2\operatorname{arccot}p < \theta < 2\operatorname{arccot}q$ 时，才有 $q < \cot\frac{\theta}{2} < p$．于是，根据 θ 的分布一致性，这件事的发生概率是 $\gamma=\frac{\operatorname{arccot}q-\operatorname{arccot}p}{\pi}$．在这种情况下，根据你猜大小的准则，不管你选的是哪一个数，你总将猜对．在其他情况下，θ 要么既大于 p 也大于 q，要么既小于 p 也小于 q，你猜对的概率是一半，因为你以同样的可能性选 p 和 q．这样，你猜对的概率就是 $\frac{1}{2}(1-\gamma)+\gamma=\frac{1}{2}+\frac{\gamma}{2}$．这正是我们刚才所断言的．

从数学上说，这样的证明无懈可击，但是这引出了关于概率论在决策上的应用方面的一些令人感兴趣的哲学问题．例如，假如你没有指针式转盘，但戴着一只手表，不是那种数字式的，

并以分针和时针之间的夹角为 θ. 从任何意义上说,这是一种随机化吗? 如果不是,为什么不是? 假如盒子里放的不是数,而是不同重量的石头. 你有一架天平,因此你能比较重量的大小,但不能称出具体重量. 你挑选一块石头,然后必须猜一下另一块石头比你那块是重还是轻. 你又能做些什么呢? 这可能导致人们认为上述随机化的方法只有在进行数量的比较才有效,也就是意味着人们必须将所选的对象联系上数字. 但是这也不一定对. 例如,假如每个盒子里放着一角馅饼,那么你只要有一个指针式转盘就够了. 可按指针间的夹角是大于还是小于你所选的那角馅饼的尖角的角度来猜另一角馅饼是比你的大还是小. 这可以用目测直接比较. 这里并不需要量角器,也根本没有涉及数字.

4. 另一个数 这次盒子里的整数是正整数,而且是相邻的. 每位局中人选一个,而且必须按以下的程序得知对方的数字. 局中人每人备有一张空白的卡片和一支铅笔. 如果在某一时刻有一位局中人知道了对方的数字,他就把它写在手中的空白卡片上,并赢得这场游戏. 如果局中人双方都不知道对方的数字,他们就交换手中的空白卡片,游戏从头开始. 我们的断言是:对于洞察能力较强的局中人,这场游戏一定会结束. 更准确地,我们把它表达为一个定理:

定理 *如果这两个数为 n 和 $n+1$,那么经过 $n-1$ 次交换后,拿了 n 的那位局中人将获胜.*

证明是用归纳法. 如果 $n=1$,那么拿了1的局中人就知道对方的数字是2,于是游戏结束,没有进行交换. 现在假设定理的结论对从1到 n 的正整数都成立,并假设现在较小的数字是 $n+1$,那么拿了这个数字的局中人将知道,如果对方拿的是 n,则对方在第 $n-1$ 次交换后就会结束游戏(归纳法假设). 而对方现在并没有这样做,因此这次交换后他就知道对方一定是拿了

$n+2$,于是他赢了.

导致这个悖论的关键在于:如果拿的两个数字比方说是72和73,那么双方都不知道对方的数字,而且双方都明白对方不知道自己的数字.于是他们无疑知道这游戏的第一步将是交换卡片.当这件事实际上做过之后他们显然什么新的信息也没有得到.然而,既然这游戏现在向结束走近了一步,就肯定要有某[9]些事情发生变化.那是什么呢[1]?

我们都会犯错误

人人都会犯错误,我们中有些人犯的错误比别人更多,因此明白甚至伟大的数学家偶尔也会发表错误的结果,可能会让人得到一些安慰.我们曾经请读者提供这种现象的例子.这里所谓错误,并不意味着仅仅是一个证明中的一个漏洞,也不是指那种意思是这而结果写了那的情况,而是指明明白白的错误断言.在高斯(C.F.Gauss)的著作中会有这样的例子吗?我不知道.鲁滨逊曾要我注意闵可夫斯基(H.Minkowski)的一个失误——他声称一个四面体的差集[2]是一个八面体.闵可夫斯基在他于1906年发表的论文《全等立体的最紧密格子式堆积》(Dichteste gitterförmige Largerung kongruenter Körper,载《格廷根皇家科学会通讯》(*Nachrichten der König. Gesellschaft der Wissenschaften zu Göttingen*) 5(1904)[3],p.311—p.355)中写道:

> 例如,如果 K 是一个四面体,那么 $\frac{1}{2}(K+K')$ 就变成一个八面体,它的各个面同那个四面体的面平行.

① 详情请参见本书第15章.——译注

② 设 A, B 是一个定义了减法的集合的两个子集,则它们的差集 $A-B$ 就是 $\{a-b: a\in A, b\in B\}$. 这里所说的一个四面体的差集,即这个四面体与其自身的差集.其中的减法即空间中点的位置向量的减法.——译注

③ 1906年的文章何以发表在1904年的杂志上?但原文如此.——译注

这里 K' 是"K 关于点 O 的对称图形".

这是一种稀奇的错误,因为所涉及的多面体显然有 12 个顶点而非 6 个顶点,即所有的 $a_i - a_j, i \neq j$[①],其中 a_i 是那个四面体的顶点.事实上,那个多面体应是一个立方体的 12 条棱的中点的凸包[②],因此它有 8 个三角形面和 6 个正方形面.

萨洛斯[③]的两个贡献

第一个如下.

"This computer-generated sentence contains two hundred forty-seven letters; four a's, one b, four c's, five d's, forty-four e's, nine f's, three g's, seven h's, eleven i's, one j, one k, three l's, two m's, twenty-nine n's, nineteen o's, two p's, one q, fourteen r's, thirty-one s's, twenty-five t's, seven u's, eight v's, seven w's, two x's, six y's, and one z."[④]

---

① 原文如此.似有误.应为 $\frac{1}{2}(a_i - a_j), i \neq j$. 这里的有关概念,有兴趣的读者可参见《凸图形与凸多面体》(杨之,劳格编译,北京科学技术出版社,1987 年版).这是一本非常精彩的小册子.—— 译注

② 这里即为包含这些中点的最小凸多面体.—— 译注

③ 萨洛斯(Lee Sallows),荷兰梅奈亨公教大学认知和信息学院博士,在下文所介绍的自描述语句以及逻辑趣题和幻方等方面颇有研究.——译注

④ 这句话译成中文就是:"这个由计算机生成的句子包含 247 个字母:4 个 a,1 个 b,4 个 c,5 个 d,44 个 e,9 个 f,3 个 g,7 个 h,11 个 i,1 个 j,1 个 k,3 个 l,2 个 m,29 个 n,19 个 o,2 个 p,1 个 q,14 个 r,31 个 s,25 个 t,7 个 u,8 个 v,7 个 w,2 个 x,6 个 y 和 1 个 z."句中的这些描述是针对这个英文句子本身而言的,这就是所谓"自描述语句"(self-descriptive sentence).在下文介绍的循环迭代算法中,又称"自描述子"(self-descriptor).把自描述语句翻译成其他语言后所描述的内容是否仍然正确,是一个很有趣的问题.萨洛斯就经常寻找从英文翻译成荷兰文后描述仍正确的自描述语句.但在中文里谈不上字母及各个字母在句子中的个数,因此上述句子译成中文后无论如何也不会成为自描述语句.在逻辑上自描述语句通常导致一种"语义悖论",最著名的语义悖论有"说谎者悖论".可参见《悖论、谬误、诡辩》(余式厚,汤军编著,浙江人民出版社,1988 年版)—— 译注

在你检验了这个句子的正确性后，你可能很想知道计算机是怎样生成这个句子的．这里是萨洛斯对这个过程的描述．

“生成上述句子的算法执行了一个迭代函数．从一个类似的语句出发，但其中说到的各个字母个数是随意写上的．现在确定出各个字母在这个语句中的真正个数，然后用以替换掉那些随意写上的字母个数．这个新的语句就成为第二次迭代的自变量，依此类推．结果产生了一系列趋向目标的近似语句．我喜欢把这个过程描绘成一台把句子拿来作为输入并把产生的句子作为输出的机器，输出的句子通过一个反馈回路返回输入端．这就使人比较容易地看出，自描述语句实际上是一种能搅乱机器的正常运行以让自己不断复制的病毒．这样一种句子只要在输入端出现一次，就会触发一个周期为 1 的闭环，于是它就被 ad nauseam①吐了出来（假定你明白我的意思）．唯一的麻烦在于，还有其他一些病毒，它们可能抢先感染了机器！这就是那些具有较长周期的句子链，遇上它们中的任何一个，机器就会很容易地陷入圈套，从而再也不能收敛到一个自描述子．我们怎样才能使机器具有免疫力以抵制这些插进来制造循环的家伙呢？

我的回答是：把机器作些修改，通过不重复的执行方式，打乱可能产生的循环．在句子每次通过机器时，我不是把所有的字母个数都作纠正，而是每次随机选择一个字母，只对这个字母的出现个数作纠正．现在，在如此不规则的替换之下，任何循环圈都无法生存．当然有一种特殊的情况除外，这时所有的字母个数都保持不变，因为它们都已经是正确的了：这就是一个自描述子！［注意这完全类似于一个神经网络在向一个稳定解状态逼近时，为避免被锁在伪解状态而采用的‘抖动’方式②．］事实

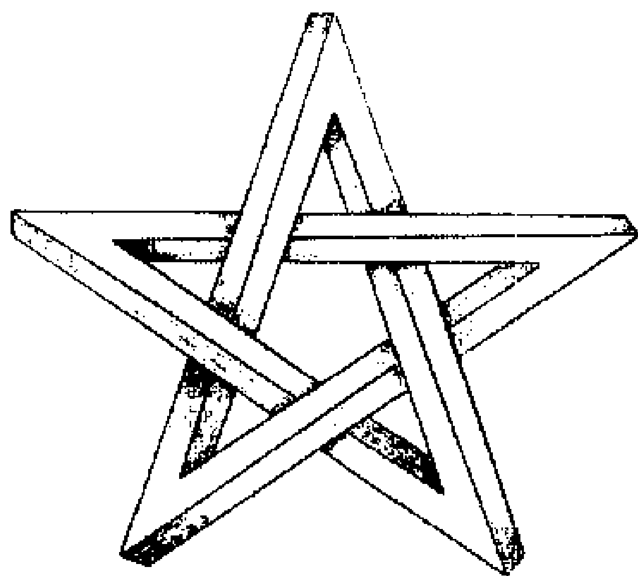
[10]

① 拉丁语，意为“令人作呕地”、“令人讨厌地”．——译注

② 这是指模拟退火法．有兴趣的读者可参见《神经网络计算》（焦李成编著，西安电子科技大学出版社，1993 年版）．——译注

上，'随机'选择并不一定要真正的随机，只要选择模式本身的重复周期比任何一个会使机器陷入困境的可能循环长就可以了．因此，任何常见的伪随机数发生器都十分适用．在正常情况下，几百万次的迭代（突变）就足以演化（自然选择）出一个活生生的解（病毒），条件是确实存在一个解．如果不行，我们可以把原来的语句作些修改再试．"

萨洛斯的第二个贡献既不包含字母也不包含数字，它如下所示，没有任何说明①．



[11]

[12]

① 把这种怪圈式的图形予以艺术化表现的人首推当代荷兰版画家艾舍尔（Maurits Cornelis Escher）．关于这类图形所蕴涵的数学哲学思想，建议读者参见《哥德尔、艾舍尔、巴赫——集异璧之大成》（侯世达著，郭维德等译，商务印书馆，1996年版）．这是一部堪称旷世奇书的著作．——译注

第3章 历史上的猜想 再说序列之谜

猜 想

近年来的数学成就中，令人瞩目的是一些长期悬而未决的著名猜想被解决了：德利涅(P. Deligne)证明了韦伊猜想[①]，德·布朗热(L. de Branges)证明了比勃巴赫猜想[②]，法尔廷斯(G. Falt-

---

① 这个由法国数学家韦伊(A. Weil)于1949年提出的猜想，由于涉及一些代数几何与代数数论的概念，看来很难用通俗简略的语言向非数学专业的读者说清楚．粗略地说，它揭示了有限域上代数方程的解的个数与由这个方程的复数解所组成的几何流形的拓扑性质之间的深刻联系．依译者之见，对这个猜想的一个比较通俗而又不失准确的叙述当推《菲尔兹奖获得者传》(胡作玄、赵斌编写，湖南科学技术出版社，1984年版)中关于德利涅(该书译为德林)的部分．一个更为通俗且反映了这个猜想关键所在的叙述可见《二十世纪数学史话》(张奠宙、赵斌编著，知识出版社，1984年版)中的"二十四、宝石集锦——30年来'难题与猜想'的进展情况"．这个猜想于1973年由比利时数学家德利涅完整地解决(他因此而获得了1978年的菲尔兹奖)．——译注

② 这个由德国数学家比勃巴赫(L. Bieberbach)于1916年提出的猜想是说：对于一个在单位圆内单叶解析的函数 $f(z) = z + a_2z^2 + a_3z^3\cdots + a_nz^n + \cdots$，有 $|a_n| \leqslant n\ (n = 2,3,\cdots)$，且仅当 $f(z) = \frac{z}{(1 - e^{i\varphi}z)^2}$ 时才有 $|a_n| = n$．该猜想于1984年由美国数学家德·布朗热解决．但他的成果初不为美国数学界所接受．后他借到苏联访问的机会就此成果作了几次演讲，受到苏联数学家重视．经国际函数论界权威学者反复仔细审查，终被承认．该猜想被彻底解决前的最好结果 $|a_n| \leqslant 1.0643n$ 由译者于1983年所得．详情可参见《现代数学新进展》(吴文俊主编，安徽科学技术出版社，1988年版)中沈燮昌教授的文章．——译注

ings)证明了莫德尔猜想①. 另一方面,费马问题②③、黎曼假设④,以及所谓的庞加莱猜想⑤仍未解决,虽然不时有人宣称这些问题也已经被解决. 无论怎么说,关于其中一些问题的起源,有几件零散的事实看来该及时地提出来. 本节下面的大部分内容,除了下面这一段,都由孟加拉国达卡大学的乔德赫里(M.R. Choudhury)教授提供,在此谨表谢忱.

---

① 这个由英国数学家莫德尔(L. J. Mordell)于1922年提出的猜想是说:设 $F(x,y)$ 为有理系数多项式,且代数曲线 $F(x,y)=0$ 的亏格不小于2,则方程 $F(x,y)=0$ 只有有限多个有理解。这里"亏格"可粗略地理解为方程 $F(x,y)=0$ 的复数解所组成的几何流形(即"代数曲线")上"洞"的个数. 该猜想由联邦德国数学家法尔廷斯于1983年解决(他因此而获得了1986年的菲尔兹奖). ——译注

② 现在不再是个猜想了! ——原注

③ 这个由著名法国数学家费马(P.de Fermat)于1637年随手写在一本书空白处的结论,终于由英国数学家怀尔斯(A. Wiles)于1994年证明(他因此而获得了1996年的沃尔夫数学奖). 详情可参见本译丛中《数学:新的黄金时代》的第8章和《费马大定理》(西蒙·辛格著,薛密译,上海译文出版社,1998年版). ——译注

④ 这个由著名德国数学家黎曼(B. Riemann)于1859年提出的假设内容如下:设复变函数 $\zeta(s)=\frac{1}{1^s}+\frac{1}{2^s}+\frac{1}{3^s}+\cdots$,其中 s 为复数且 $\mathrm{Re}s>1$. 用解析延拓的方法可将 $\zeta(s)$ 的定义域开拓到整个复平面. 这时,$\zeta(s)$ 有一系列一级零点:$s=-2,-4,\cdots,-2n,\cdots$,它们被称为"平凡零点". 除此之外,$\zeta(s)$ 的零点都被称为"非平凡零点". 黎曼假设:$\zeta(s)$ 的所有非平凡零点都位于直线 $\mathrm{Re}s=1/2$ 上. 黎曼假设对于数论和函数论具有极其重大的意义,在黎曼假设成立的前提下获得的重要结论一般都被认为是有价值的. 近一个半世纪以来,虽然获得了不少支持黎曼假设的成果,但黎曼假设本身既没有被证明也没有被否定. ——译注

⑤ 这个以著名法国数学家庞加莱(J.-H. Poincaré)的名字命名的猜想将在文中叙述. 这里要补充的是,后来有人在此基础上提出了所谓广义庞加莱猜想:当 $n\geq 4$ 时,单连通的 n 维闭流形若与 n 维球面有相同的同调群,则必与 n 维球面同胚. 对庞加莱猜想及广义庞加莱猜想的一个避开了"单连通"、"闭流形"、"同调群"、"同胚",以及文中的"紧流形"、"基本群"等拓扑学概念的直观描述,由本译丛中《数学:新的黄金时代》的第10章给出. 关于广义庞加莱猜想,$n\geq 5$ 的情况已由美国数学家斯梅尔(S. Smale)于1960年证明(他因此而获得了1966年的菲尔兹奖),而 $n=4$ 的情况则由美国数学家弗里德曼(M. Freedman)于1981年证明(他因此而获得了1986年的菲尔兹奖). 现在只剩下庞加莱猜想本身未获解决. ——译注

通常说到庞加莱猜想，是指这样的断言：只有单连通的 3 维紧流形才是 3 维球面[①]. 然而正如有人指出的那样（例如可参见斯梅尔发表在《数学信使》（*Mathematical Intelligencer*, Vol. 12, [13] No. 2, 1990）上的文章），庞加莱从来没有把它作为一个猜想提出来. 他在《关于位置分析的第五个补充》（Cinquième complément à l'analysis situs, 见《著作第六卷》（*Œuvres* Ⅵ）, Gauthier-Villars, Paris (1953), p. 498）中写道：

> Il resterait une question à traiter: Est-il possible que le groupe fondamental de V se réduise à la substitution identique, et que pourtant V ne soit pas simplement connexe?
>
> 还剩下一个问题要处理. 即使 V 不是单连通的，它的基本群是不是也能约化成单位元群？

接着有一段文字，其中用这篇文章引入的一些概念把这个问题重新叙述了一遍. 然后以一段只有一行的文字作为结束：“Mais cette question nous entranerait trop loin.”它可粗略地翻译为：“但这个问题将把我们带往遥不可及的远方.”可见这个问题既不是作为一个猜想，也不是作为一个公开让人们研究的问题而提出的. 庞加莱对答案连猜都没有猜.

与此相反，韦伊以绝对明确的态度相信他那个还没有证明的结论是成立的. 他在《有限域上方程的解的个数》（Numbers of solutions of equations over finite fields, 载《美国数学会通报》（*Bulletin of the American Mathematical Society*）55(1949), p. 498）中这样写道：

> 这个例子，以及其他我们不能在此讨论的例子，看来给予下述猜想以某种支持. 这个猜想已知对代数曲

① 这里的“是”，意思是“与……同胚”. 在拓扑学中，均把相互同胚的流形视作同一. ——译注

线是成立的，但是至今我还不能证明它对更高维的代数簇[①]也成立．

正如我们现在从德利涅的工作所知道的，韦伊猜想结果被证明是正确的．于是，看一看韦伊就猜想这个主题所作的一般性论述(见《关于数论的两个演讲，过去与现在》(Two lectures on number theory, past and present，载《数学教学》(*L'Enseignement mathématique*) (2)，20(1974)，p.87—p.110)，是一件很有意思的事：

这里我可以指出，在过去的日子里，当我们使用“假设”或“猜想”(德文作 Vermutung)这个词时，并不是简单地把它作为一种一厢情愿的想法．如今，这两方面经常混淆．例如，关于丢番图方程的所谓“莫德尔猜想”是说，一条亏格至少为2的有理系数代数曲线最多有有限多个有理点．如果真是这样，那是很好的．我宁愿打赌说它成立而不愿否定它．但这只不过是一个一厢情愿的想法，因为既没有丝毫证据来支持它，也没什么来否定它．

因此，庞加莱显然没有作出他的猜想，而韦伊则确实作出了猜想．我们让读者来决定莫德尔的下列陈述(见《关于三次和四次不定方程的有理解》(On the rational solutions of the indeterminate equations of the third and fourth degree，载《剑桥哲学会会刊》(*Proceedings Cambridge Philosophical Society*) 21(1922/23)，p.191—p.192)是一种猜想，还是一种一厢情愿的想法：

总之，我可以注意到，前面的工作向我提示了下列关于不定方程的陈述的正确性，然而，其中任何一个我

① 粗略地说，代数簇是指代数方程的复数解所组成的几何流形．1维的代数簇又称“代数曲线”；2维的代数簇又称“代数曲面”．——译注

都不能证明……

(3) 方程

$$ax^6 + bx^5y + \cdots + fxy^5 + gy^6 = z^2$$

只能被 x 和 y 的有限多个有理值所满足，并且这一结论显然可以推广到更高次的方程.

(4) 同样的定理对方程

$$ax^4 + by^4 + cz^4 + 2fy^2z^2 + 2gz^2x^2 = 2hx^2y^2 = 0$$

也成立.

(5) 同样的定理对任何亏格大于1的齐次方程，例如 $f(x,y,z) = 0$，也成立.

再说序列之谜

许多数学家已经注意到，计算机对数学的主要影响在于提出新问题，而不是在于解决老问题. 我要指出，在这些新问题中，有一些问题虽然表述起来很容易，但是或许在实际上是不可能解决的.

在那些如果没有计算机恐怕永远也不会提出来的问题中，一个最简单而且最闻名的问题是科拉茨(Lothar Collatz)的所谓 $(3n+1)$ 猜想. 令 f 是自然数集 $\mathbf{N}$ 上的一个函数，且

$$f(n) = \begin{cases} n/2, & \text{当 } n \text{ 为偶数}, \\ 3n+1, & \text{当 } n \text{ 为奇数}. \end{cases}$$

这猜想是说，对任何一个 n，都相应地存在一个 k，使得 $f^k(n)$[①] $= 1$. 或者用动力系统的语言来说，就是每个 n 的轨道[②]都包含1. 这一结论对所有的 $n < 10^9$ 都已得到验证.

容易描述以此为特殊情况的一类广义问题. 对 $\mathbf{N}$ 中的任何

---

① 这里的 $f^k(n)$，并不是指 $f(n)$ 的 k 次方，而是指用 f 对 n 进行 k 次迭代运算，即 $\underbrace{f(f(f(\cdots f}_{k个}(n)\cdots)))$. ——译注

② 所谓 n 的轨道，即序列 $f(n), f(f(n)), \cdots, f^k(n), \cdots$. ——译注

k 以及 a_i, b_i[①], $0 \leq i < k$, 令 f 为从 **N** 到 **N** 的一个函数, 其定义为

$$f(kn + i) = a_i n + b_i.$$

于是我们可以问关于点在 f 作用下的轨道问题. 例如, 它们是否都包含 1? 科拉茨问题就对应着 $k = 2, a_0 = 1, b_0 = 0, a_1 = 6, b_1 = 4$ 的情况.

关于这个广义问题的主要结果是由康韦(John H. Conway)予以证明的(见《不可预测的迭代》(Unpredictable Iterations, 载《1972年博尔德数论会议论文集》(*Proceedings Number Theory Conference, Boulder*, 1972)). 这个结果断言: 它是不可判定的. 更准确地说, 康韦证明, 即使是对所有的 b_i 都为零的情况, 也没有一个算法能判定, 对任何一个给出的 n, 其轨道是否包含着 1. 当然, 由此得不出什么关于科拉茨问题可判定性的结论, 然而它确实证明了, 虽然存在着一些有着参数 k 和 a_i 的具体问题, 它们在理论上是可以算出的, 但是对于这个问题类来说, 却是不可判定的. 康韦进一步得到了一些有趣的例子, 它们被盖伊(Richard Guy)称为"置换序列". 其中最简单的例子出现在盖伊的《数论中的未解决问题》(*Unsolved Problems in Number Theory*, Springer-Verlag, New York, 1981)[②]中, 并且由如下定义的映射 T 给出:

$$T(2n) = 3n,$$
$$T(4n+1) = 3n+1,$$
$$T(4n-1) = 3n-1.$$

15

容易看出 T 是一个一一映射, 因此所有的轨道要么是圈, 要么

---

① 一般来说, 自然数集不包括 0, 但从下文看, 这里的 b_i 可以为 0. ——译注

② 有据此书改写的中文本,《数论中的问题与结果》(曹珍富编著, 哈尔滨工业大学出版社, 1996年版). 有兴趣的读者可参阅. 但是其中把康韦的例子译为"排列序列". ——译注

是左右都趋向于无穷的双无穷序列. 人们可以容易地发现(1),(2,3),(4,6,9,7,5)和(44,66,99,74,111,83,62,93,70,105,75,59)这些圈. 不在这些圈中出现的最小的数是8,它的轨道开头是73,55,41,31,23,17[①],13,10,15,11,**8**,12,18,27,20,30,45,34,51,38,57,43,32,48,72,…. 这条轨道经过在两个方向上的几千次迭代,虽然发生了一些"擦肩而过"的现象,但毫无首尾相接的迹象. 请注意在上述序列两头的73和72. 这种现象多次发生,如153和154,161和162,500和501,790和791.

问题1　在 T 下是否有任何其他的有限轨道?

问题2　8的轨道是否有限?

后一个问题有一个令人注意的特点,即它针对的是单个序列. 这同科拉茨问题相反,后者问的是无穷多个序列的行为. 当然,说这个问题不可判定是没有意义的,但下面我将论证,它很可能是"不可证明的". 也就是说,8的轨道可能在实际上是无穷的,但对此从我们通常的公理系统出发根本不存在证明.

进一步的实验导致进一步的推测. 不在上述所有轨道中出现的最小数是14,它的轨道看来义无反顾地从两头向无穷延伸. 下一个不出现的数是40,如此等等. 用 *Mathematica* 我们发现所有1000以内的数分属54条不相交的轨道. 我们把一条轨道中的最小数称为**种子** s. $T^n(s)$ 中的元素,在 n 为正的情况下称为**向前数**,在 n 为负的情况下[②]称为**向后数**. 当然,就像我们根本无法证明这些轨道是否一个圈的一部分一样,我们也不知道它们是否不相交. 例如,可以想像,$T^n(8)$ 向前迭代到某处最终可能与 $T^m(14)$ 的一个向后数相遇. 一个粗略的统计研究表明,向前数中包含的偶数与奇数个数大致相同. 由此可以推出,约有一半的向前数能被3所整除,因为根据 T 的定义,当且仅当

① 原文缺17.——译注

② 即用 T 的逆函数进行 n 次迭代运算.——译注

一个数的前一个数为偶数时，这个数才被3所整除．另一方面，在向后数中，奇数看来以大约2对1的比例多于偶数．情况显然一定是这样，因为在向后迭代时，一个偶向后数减小到它前一个数的2/3，而一个奇向后数只增大到它前一个数的大约4/3．因此向后数中奇数一定比偶数多，因为 $T^{-n}(s)$ 作为自然数，必须保持为正．

利用手中的数据可以证明，如果除这里列出的以外还有其他任何的圈，那么它们的长度必定至少为360．而且，我们可以直观地论证，圈的存在性变得非常的“不可证明”．首先请注意，任何奇数在向前迭代一次后都（近似地）减小到原来的3/4，而任何偶数则增大到原来的3/2．因此如果一个圈有 m 个奇数和 n 个偶数，则我们必须有 $(3/4)^m(3/2)^n \approx 1$．而如果这个圈中最小的数，即种子，在1000附近，那么这个近似等式一定误差很小，也就是说，比值 m/n 一定同0.70951很接近，精确到第四位小数．如果种子大于10 000，那么最小的圈也将有665的长度，而且只能在 $m = 276$ 和 $n = 389$ 的情况下存在．发生这种情况的“机会”是多少呢？

那么，这个故事的寓意何在呢？我们都知道，根据哥德尔（K.Gödel）的工作，不管我们所据的公理系统是什么，只要用到那么一点儿数论，就存在一些真命题，对它们根本不存在什么证明[①]．然而我们还是一如既往，锲而不舍地寻找着证明，而且我们往往找到了它们．这里的原因，我想主要在于我们选来攻克的问题．但是像我们在这里讨论的问题，看来是一种特殊类型 [16] 的问题．在我看来，8的轨道为无限的“可能性”是无法抗拒的，对这个事实作出证明的“不可能性”同样也是无法抗拒的．说实在的，为什么应该有一个证明呢？

---

① 这是指著名的哥德尔不完全性定理．有兴趣的读者可参见《哥德尔不完全性定理》（朱水林编著，辽宁教育出版社，1987年版）．——译注

怀特黑德的幽默

拓扑学家J·H·C·怀特黑德(J.H.C.Whitehead)经常被问到他对他的叔叔——著名哲学家A·N·怀特黑德(Alfred North Whitehead)的工作的看法.后来,他想出了一个固定不变的回答.当被问到"你认为你叔叔的哲学怎么样"时,他就答道:"我真的对此没有很好想过——然而你认为你叔叔的哲学又怎么样呢?"[17]

[18]

第4章 保护个人隐私的协议

无条件安全的协议

过去的一个年代中，令人激动的数学进展之一是发现了所谓不可破译的公钥密码．这些密码具有这样的特征：人人都知道加密的方法，但对外人来说，破译这个密码所需要的计算量被认为超出了当前计算机的能力．在最著名的例子中，要破译这个密码，相当于要找出一个大整数（比方说一个100位整数）的所有因数．人们确信这在计算机上是不可行的．

一个不那么著名但在某种程度上有着同样风格的更新进展，是关于传递某种信息的方法，这种信息取决于其他信息，而那些其他信息则必须保密．然而，用这些方法，除了自称能洞悉他人心灵的巫师外，任何人都绝对不可能得知保密的信息．这里有一个简单的例子．有一些人，设为 $P_1, P_2, \cdots, P_n$，比方说是一个数学系的全体成员，很想知道他们的平均薪金是多少．但他们都不愿意把自己的薪金数泄露给其他任何人．怎样才能做到这一点呢？我把这个问题提给我的一些同事，但他们都不能给出答案．我在一次社交聚会上又提到了这个问题，一位读完 [19] 高中以来什么数学课程都没有进修过的妇女（而且声称她没能通过九年级的代数考试）很快就作出了如下的简单解法：P_1 随意选一个数 x，并把它告知 P_2；P_2 把自己的薪金数加上去，并把所得的和告知 P_3；P_3 也把自己的薪金数加上去，并把和告知

P_4. 依此类推. 最后 P_n 把自己的薪金数加上去并把和告知 P_1，P_1 加上自己的薪金数，再减去 x，除以 n，并把结果公布于众. 显然，除了可由平均薪金推断出的信息外，没有人能得知关于其他人薪金的任何情况.

在上述方案中，确实没有人能通过自己独立的行动发现关于其他人薪金的任何情况，但是，如果现在允许人们可以合谋，情况就会改变. 例如，如果 P_1 将 x 泄露给 P_3，那么 P_3 将得知 P_2 的薪金数. 于是，这个方案，或者如一般所称的**协议**(protocol)，被称为是 **1－保隐私的**(1-private)，但不是 **2－保隐私的**(2-private). 于是有人要问，对这个问题是否存在 2－保隐私的协议. 回答是，事实上存在着一个 n－保隐私的协议，而且这个协议也很容易描述. 一个协议被称为是 ***n*－保隐私的**(n-private)，是指这 n 个人的任何一个真子集都不能通过合谋得知关于它的补集的任何情况，除了他们知道了平均薪金后由此推断出的信息. 下面是这个协议的执行过程. 令 s_i 是 P_i 的薪金数. 每位 P_i 都选 n 个数 s_{ij}，它们只要满足加起来等于 s_i 这个条件. 然后 P_i 将每个 s_{ij} 各写在一张纸片上，并将写有 s_{ij} 的纸片交给 P_j. 现在每位 P_j 将手中纸片上的数字之和 t_j 公布于众. t_j 之和当然就是 s_i 之和，这正是人们想知道的数字. 有关情况可用矩阵 $\mathbf{S}=(s_{ij})$ 来表示，其中第 i 行的数字之和是 s_i，而第 j 列的数字之和是 t_j.

	t_1	t_2	t_{k1}	t_n
s_1	s_{11}	s_{12}	s_{1k}	s_{1n}
s_2				
s_k	s_{k1}	s_{k2}	s_{kk}	s_{kn}
			$\mathbf{S}_k$	
s_n	s_{n1}	s_{n2}		s_{nn}

要看出这个协议是 n - 保隐私的，不妨假设前 k 个参预者合谋. 那么除了右下角那个 $(n-k)\times(n-k)$ 的子矩阵 S_k 中的元素外，S 中的其他所有元素都为他们所知①，而且他们也知道 $t_{k+1},\cdots,t_n$，因此他们将知道 S_k 的所有列和. 但如果只知道一个矩阵的列和 c_j，那么这个矩阵的行和 r_i 就可以是任何仅满足条件 $\sum r_i=\sum c_j$ 的数. 因此这些合谋的参预者只能知道其他参预者的薪金之和，而这一点无论怎么说都可在他们知道了薪金总和后推断得知②.

这个关于和的协议可以用来得知其他一些信息，例如薪金的分布情况，即在每个薪金收入水平上各有多少个人，但是这些人是谁则不得而知. 要知道有多少人的薪金是 x，可以执行上述关于和的协议，不过 P_i 的秘密数取1或0：如果他的薪金数为 x，就取1；否则，就取0. 对 x 的所有值③重复执行这个协议. 一个更有效的办法是：对 P_i 来说，取 $(n+1)^{s_i}$ 为其秘密数. 用上述协议算出秘密数的总和，并且把它以 $(n+1)$ 为底数表示出来④，$(n+1)^x$ 的系数就是薪金为 x 的人的个数. 同样的技巧可以用于无记名投票. 假定有若干个候选人要竞选某个职位，将他们用1到 m 标记. 一个想投票给候选人 k 的选民就应该用秘密 [20]

---

① 原书所附的矩阵 S 示意图如上页所示，但似有误，子矩阵 S_k 不应包括 S 的第 k 列元素，特别是，不应包括 $s_{k+1,k},\cdots,s_{nk}$. ——译注

② 需要说明的是，当 $k=n-1$ 时，这 $n-1$ 个参预者可由这个协议得知那个被他们合谋暗算的人的薪金数，但这一点同样可由薪金总和推出. 由此可体会到在 n - 保隐私的定义中"除了他们知道了平均薪金后由此推断出的信息"这句话的意义. ——译注

③ 这里显然假定薪金数只取有限个值，而且可以假定只取有限个非负整数值. 这种假定是合理的，因为薪金一般都分为有限个级别，而且若取最小货币单位，薪金数即为非负整数. 虽然薪金数为零似不太有现实意义，但这在数学上多有方便之处. ——译注

④ 即表示为一个 $n+1$ 进制数，下文 $(n+1)^x$ 的系数即这个 $n+1$ 进制数的右数第 $x+1$ 位数字. ——译注

数$(n+1)^k$. 于是,这个关于和的协议将会给出各候选人的得票统计.

其他函数的情况怎样?例如,想要知道的是最高薪金而不是薪金总和,那该怎么办?如果已知薪金数的一个上界$\bar{s}$,那么自然就会想到下述的程序.用这个关于和的协议算出有多少人的薪金为$\bar{s}$. 如果答案为零,就再试$\bar{s}-1$,如此进行下去,直到算出的和为正数.这里的问题是人们知道得太多了.人们不但知道了最高薪金,而且还知道了拿最高薪金的人的个数.是不是可以让人们仅得知最高薪金而对其他情况一无所知?同样地,是不是可以让人们在某次选举中仅得知哪位(哪些)候选人当选,但是对选票的分布情况一无所知?还有一个简单的算术问题:既然n个数的和可以用n-保隐私的协议算出,那么对n个数的积,情况又如何?

所有这些问题的答案不但是同样的,而且十分令人意外.当且仅当t小于$n/2$时,它们都存在t-保隐私的协议.这种协议可称为**少数人-保隐私的**(minority-private).少数人-保隐私的协议的存在性是由本-奥(M.Ben-Or)、戈德瓦泽(S.Goldwasser)和威格德森(A.Wigderson)[2]证明的,它也为肖姆(D.Chaum)、克雷波(C.Crépeau)和达姆加德(I.Damgard)[3]独立地证明.给定秘密数$s_1,s_2,\cdots,s_n$,它们可以取自某个有限的数值集合,则s_i的任何函数都可以用少数人-保隐私的协议算出.只要考虑在一个充分大的有限域上的函数就可以了.对于乘法来说,已经给出了一个少数人-保隐私的协议,它在某种程度上远比那个关于加法的协议来得复杂.(迄今为止我们描述的所有内容可以为一名合格的七年级学生所理解.关于乘法的这个协议则大约在大学抽象代数课程的水平上.)一旦解决了加法和乘法的情况,我们就解决了多项式的情况,从而一个有限域上所有函数的情况都得以解决.通过适当的编码,人们感兴趣的大多数问题都可以变换为计算一个从整数到整数的函数的问题,虽然

在有些情况中,这一点并不是一目了然的,例如在人们只想知道哪位候选人赢得选举的无记名投票问题中.

或许比条件 $t < n/2$ 的充分性更为令人惊奇的是它的必要性. 这意味着,比方说在计算 n 个秘密数之积的问题中,如果这 n 个参预者中有半数或半数以上的人决定合谋,那么没有一个协议能做到保密. 事实上,唯一能够用**多数人-保隐私的**(majority-private)协议计算的函数,本质上就是仅用关于和的协议算得的函数. 这一点首先由肖尔(B. Chor)和库什列维奇(E. Kushilivitz)[4]对布尔函数得到证明,后来又由比弗(Donald Beaver)[1]对一般的整数值函数得到证明. 请注意到现在为止我们还没有说过协议到底是什么,我们只不过简单地展示了一些例子. 只要人们证明的是存在性定理,这样做一切顺利. 类似地,要证明对三次和四次多项式存在一个求根"公式",人们只要简单地把公式写出来并检验它们确实有效就行了. 但另一方面,要证明对更高次的多项式来说这类表达式的**不存在性**,就有必要把问题予以严格的正式化. 同样,要证明多数人-保隐私的协议的不存在性,人们必须有关于协议和隐私的精确定义,然后发展出处理这些概念所必需的理论,而有关的论证比起关于存在性的论证来说,远为复杂棘手.

[21]

作为比弗结果的一个特殊情况,我们看到,当只有两个人的时候,如果他们的秘密不被泄露,那么对他们的有关情况,例如他们是否具有相同的秘密数,本质上什么都不能得知. 另一方面,根据存在性定理,我们知道,如果有一位第三者 P_3 加入进来,而且他能够给出和接收消息,那么 P_1 和 P_2 就可以用一个1-保隐私的协议得知他们是否具有相同的秘密数,而 P_3 则不知道是相同还是不同.

除了这里提到的以外,这个理论还有大量的内容. 例如,如果人们不要求无条件的安全性,而只是要求在本章第一段文字中所描述的意义上的"不可破译性",那么已经证明任何函数都

可以用 n - 保隐私的协议计算，包括只有两个人的情况．例如，在姚期智（A. C. Yao）[5]提出的所谓“百万富翁问题”[①]中，P_1 和 P_2 就可以得知他们两人中谁的薪金更高，但是对其他情况则一无所知．

作为总结，让我们回到七年级的水平，来描述一个计算最高薪金的 1 - 保隐私的协议．为此我们引进一位外来者 P_0，他选择了某个秘密数 x_0．接下来的规则是：如果 P_i 的薪金是 s（上界），那么他就任意取某个正数 x_i 为秘密数；如果不是，他的秘密数就是 0．现在执行关于和的协议．如果和不是 x_0，那么 P_0 宣布 s 就是最高薪金．如果和是 x_0，就用 $s-1$ 代替 s，重新执行协议．如此进行下去，直到最高薪金被求出．请注意引进 P_0 是必要的，因为如果没有他，其他人进行到某一步如果得出的和为 x_i，则 P_i 就知道自己是唯一拿这份最高薪金的人．同样，这个带有 P_0 的协议仅仅是 1 - 保隐私的，因为如果 P_0 与一个拿最高薪金的人串通一气，那么他们两人就会知道是否还有其他人也拿这最高薪金．

我要对美国电话电报公司（AT&T）的比弗表示感谢，我在这里引用的材料有许多是他提供的；我还要对加利福尼亚大学伯克利校区的赫希（Michael Hirsch）表示感谢，是他使我对这个有趣的主题引起了注意．看来天上人间的数学门类比你那布尔巴基[②]的皇皇巨著所梦想的还要多．

---

① 百万富翁问题是说：有两个百万富翁，要比一比他们中谁更富有，但又都不愿意把自己的财产数量暴露出来，要求设计一个协议，能够得出正确的比较结果．这样的一个协议可参见《计算机密码学——计算机网络中的数据保密与安全（第2版）》（卢开澄编著，清华大学出版社，1998年版）．这个协议中使用了不可破译的公钥密码．由此可体会到“无条件的安全性”与“不可破译性”的区别：如果破译了其中的公钥密码，百万富翁的财产数量即可得知．——译注

② 布尔巴基（Bourbaki）是20世纪30年代开始在法国形成的一个数学学派，其思想主要反映在多卷集《数学原理》（Élémernt de Mathematiques）中．详请可参见《布尔巴基学派的兴衰》（胡作玄编著，知识出版社（上海），1984年版）．——译注

参考文献

1. D. Beaver, Perfect privacy for two-party protocols, *Harvard Tech. Report* TR-11-89, Aiken Computer Laboratory.
2. M. Ben-Or, S. Goldwasser, and A. Wigderson, Completeness theorems for non-crytographic fault tolerant distributed computations, *Proceedings 20th STOC*, 1988, 1—10.
3. D. Chaum, C. Crépeau, and I. Damgard, Multiparty unconditionally secure protocols, *Proceedings 20th STOC*, 1988, 11—19.
4. B. Chor and E. Kushilevitz, A zero-one law for Boolean privacy, *Proceedings 21st STOC*, 1989, 62—72.
5. A. C. Yao, Protocols for secure computation, *Proceedings FOCS*, 1982, 160—164.

关于索莫斯序列的最新报道

在第1章中,我们描述了由一种简单的递归关系定义的一些序列,这种递归关系看来以不能解释的理由总是产生整数项.这些序列最初是由索莫斯引进的,它们可以如下描述:给定一个 [22] 整数 $k \geq 4$,一个**索莫斯 k 序列**由递归关系

$$a_n a_{n-k} = x_1 a_{n-1} a_{n-k+1} + x_2 a_{n-2} a_{n-k+2} + \cdots + x_r a_{n-r} a_{n-k+r} \quad (1)$$

所刻画,其中 $r = [k/2]$,x_i 是给定的整数.

既然在(1)中 a_n 用它前面的 k 个项来定义,那么人们必须为 $a_0, a_1, \cdots, a_{k-1}$ 选定**初始值**.

当我只是简单地说"索莫斯 k"时,我的意思是指一个 x_i 及初始的a_i都为1的索莫斯 k 序列. 人们先是从数值上观察到,后来又证明了,索莫斯4,5,6,7总是具有整数项,而索莫斯8,9并非如此,而且推测其余的索莫斯 k 也并非如此. 但这仍然留下了一个二重无穷的索莫斯序列族,它们看来都具有整数项,虽然这一点还未得到证明.

为索莫斯现象所激发，斯科特发现初始值为 1 且有下述递归关系的序列总是具有整数项：

$$a_n a_{n-k} = a_{n-1}^2 + a_{n-2}^2 + \cdots + a_{n-k+1}^2, \tag{2}$$

以及，对 k 为奇数，

$$a_n a_{n-k} = a_{n-1} a_{n-2} + a_{n-3} a_{n-4} + \cdots + a_{n-k+2} a_{n-k+1}. \tag{3}$$

这个总是产生整数的性质也为

$$a_n a_{n-k} = a_{n-1} a_{n-2} + a_{n-2} a_{n-3} + \cdots + a_{n-k+2} a_{n-k+1} \tag{4}$$

所具有.

关于(2)，(3) 和(4) 的整数性的证明现在已由鲁滨逊发现，他在某一点上得到了希克森的帮助. 这些证明十分初等，所以我们将介绍关于(2) 的证明. 对(3) 和(4) 的论证过程与此相似. 它们都涉及寻找在这些递归关系下成为**不变量**的有理函数. 对于(2)，我们定义一个新序列如下：对 $n \geqslant k$，

$$b_n = \frac{a_n + a_{n-k}}{a_{n-1} a_{n-2} \cdots a_{n-k+1}}. \tag{5}$$

我们宣称，b_n 是常数. 也就是说，$b_{n+1} = b_n$. 要证明这一点，请注意

$$a_n(a_n + a_{n-k}) = (a_{n+1} + a_{n-k+1}) a_{n-k+1}. \tag{6}$$

这是因为，根据(2)，(6) 的两边都等于 $a_n^2 + a_{n-1}^2 + \cdots + a_{n-k+1}^2$. 用 $a_n a_{n-1} \cdots a_{n-k+1}$ 同除两边，即给出 $b_n = b_{n+1}$. 但由初始条件，有 $b_k = k$，于是 $b_n = k$. 因此由(5)，

$$a_n = k(a_{n-1} a_{n-2} \cdots a_{n-k+1}) - a_{n-k}. \tag{7}$$

这就给出了关于 a_n 的一个新的递归关系，其中右边是 a_i 的一个多项式(而不是一个有理函数). 于是，整数性唾手可得，同样可得的是这样的事实：对任何 m、这个序列以模 m 取剩余后将具有**周期性**.

然而，虽然有些问题现已解决，但鲁滨逊的进一步数值探索揭示了索莫斯序列的一大批构造性质，它们有一些是数论性质，

另一些则是解析性质. 由于这些结果将在其他地方发表,我只提其中少数几个. 首先,对任何 m,所有给出整数的索莫斯序列 [23] 以模 m 取剩余后都具有周期性. 鲁滨逊对索莫斯 4 和 5 证明了这一点,但对索莫斯 6 和 7 则还没有. 对于索莫斯 4 和 5,它们的周期作为 m 的函数是不可预测的,但观察到的下述关系令人吃惊:对于除了 2 以外的任何 m,以模 m^k 取剩余后的周期等于 m_{k-1} 乘上以模 m 取剩余后的周期. 对于 2 来说,则成立着一个在某种意义上更为复杂的模式. 鲁滨逊还研究了哪些素数能整除序列中的不同项,他发现对索莫斯 4 和 5 来说,能被一个给定素数整除的项呈等距分布(但索莫斯 6 和 7 则不是). 例如,在索莫斯 4 中,每隔 4 项出现一个偶数,每隔 6 项出现一个能被 11 整除的数,但没有一项能被 5 整除;而在索莫斯 5 中,每隔 9 项出现一个能被 5 整除的数,但没有一项能被 7 整除.

在一个不同的方向上,鲁滨逊研究了 k 和初始值均为任意的序列的解析性质,他发现在所有检测过的情况中,存在着(唯一的)常数 C 和 D,使得

$$a_n = C^{(n-D)^2}\phi(n).$$

其中 $\phi(n)$ 是一个振荡函数,它具有一个定义良好的周期. 常数 C 和 D 依赖于初始值和 k,但是依赖的方式显然是不可预测的. 在得知了鲁滨逊的数据后,克利福德 · 加德纳(Clifford Gardner)成功地发现了索莫斯 4 和 5 的用雅可比椭圆函数表达的精确公式. 因此在某种意义上说,这个问题正要画成一个完美的圆圈,因为当初索莫斯就是在研究椭圆函数的性质时发现了他的序列并在后来逐渐得知这里描述的一些现象的.

对于一位行外人来说,索莫斯序列的解析性质与数论性质之间似乎没有什么联系,但或许代数数论家们将能建立一种联系,他们习惯于做这样的事. 无论如何,看到这些序列由数值探索揭示出越来越多的性质,是一件令人心驰神往的事.

一个真实的故事

从前有一位小姑娘，名叫克拉拉，她只有三岁，刚学会数数．她能够告诉你起居室里有几张椅子，前门口有几级台阶．一天，她爸爸决定测验她一下．“你瞧，”他说，“我为你买来了这四支冰棍．”但他只递给她三支．克拉拉接过冰棍，照例数了起来，“一，二，四．”于是她显出有点迷惑的神色，问道，“这第三支在哪儿呢？”[24]

第5章 出人意料的洗牌

精心洗牌切牌，结果混沌一片

想像(如果你能想像的话)一副有着可数无穷多张牌的纸牌．每张牌都用一个不同的自然数标记．起初，这些牌按自然数顺序叠置，1号牌在最上面，正面朝下地放在一张(有限的)桌子上．

定义 一次完全的 n - 洗牌(perfect n-shuffle)是指取这副牌的上面 n 张牌，并把它们一张隔一张地插到紧接在其下的 n 张牌中去．例如，如果人们对这副牌在它初始顺序下执行一次5 - 洗牌，结果顺序将是

$$6,1,7,2,8,3,9,4,10,5,11,12,13,\cdots$$

现在考虑执行一系列洗牌，首先是一次1 - 洗牌，然后一次2 - 洗牌，再然后一次3 - 洗牌，如此等等．

猜想 在这一系列洗牌的过程中，每张牌都将无穷多次地来到这副牌的最上面．

这个猜想是对盖伊的一个猜想稍作修改而得来的，这一点将在下面说明．因此从现在开始，我就把这个猜想称作盖伊猜想．

25

这里是前八次洗牌给出的顺序：

0　1,2,3,…

1　2,1,3,4,5,…

2　3,2,4,1,5,6,…

3　1,3,5,2,6,4,7,8,9,…

4　6,1,4,3,7,5,8,2,9,10,…

5　5,6,8,1,2,4,9,3,10,7,11,12,…

6　9,5,3,6,10,8,7,1,11,2,12,4,13,14,…

7　1,9,11,5,2,3,12,6,4,10,13,8,14,7,15,16,…

8　4,1,10,9,13,11,8,5,14,2,7,3,15,12,16,6,17,18,…

请注意到这一步,从 1 号牌到 6 号牌都已经成功地到过这副牌的最上面,1 号牌已经到过三次.然而,7 号牌直到第 78 次洗牌时才到那儿.这就给出了这种捉摸不定的行为的第一个征兆.这种行为马上就予以描述.

回到盖伊的猜想,存在什么根据可以相信它可能是成立的?对此有两个论据,第一个是"经验的",第二个是"概率的"(这里用了引号,其意思是,从严格的意义上说,我们将要讲述的论据全是胡说八道).我们先考虑概率的论据,这是鲁滨逊提出的.首先请注意,c 号牌在第$[c/2]$次洗牌前总是呆在它的老位置上,此后它就可能以看起来骚动不安的方式上窜下跳;但在第 n 次洗牌时,它绝不可能到达一个比从最上面数下来第 $2n$ 张牌更远的地方,因此在第 n 次洗牌时,c 的位置号是从 1 到 $2n$ 当中的某个数字.现在假定(这里就是胡说八道开始插入的地方) c 的位置号是随机的,于是它不在这副牌最上面的概率就是 $1-1/(2n)$,而且(更是胡说八道)假定在接连的洗牌中位置号是独立的,于是我们知道,c 永远到不了这副牌最上面的概率就是

$$\prod_{n=1}^{\infty}\left(1-\frac{1}{2n}\right).$$

由于调和级数的发散性,这等于零.

至于"经验的证据",阿德勒(Ilan Adler)对 1 号到 5000 号的

所有牌都进行了检测，并列出了每张牌到达这副牌最上面时洗牌的次数，结果十分有趣．在 39 号牌之前，情况相对来说比较平静，39 号牌经过 13 932 次洗牌后到达了最上面．此后情况又归于安定，虽然 43 号牌需要 30 452 次洗牌．但是接下来发生了一次大暴乱，53 号牌只要洗牌 30 次，而 54 号牌居然野蛮地横冲直撞（54 号牌，你跑到哪里去了？），最后在第252 992 198次洗牌时才把它制服在最上面．

所有这些数据，是在一个 NeXT 工作站上花了大约 80 个小时的计算机时间才得到的．但是其中的大多数时间都用来对付三个"巨怪"，4546，3729，和当今的世界冠军，3464．它们分别需要2 263 846 432，15 009 146 841和21 879 255 397次洗牌．当然，从那荒谬的概率论据出发，这种行为倒是在人们的预料之中．一张牌离开最上面的距离越长，它继续处于这种状态的时间可能也越长；也就是说，在第一百万次洗牌时"选中"一个 1 的机会是一百万分之一．

既然完全的洗牌表现得如此桀骜不驯，或许人们应该把眼光放在较为简单的事情上．我们可以先试一试用切牌来代替洗牌．传统的切牌就是取最上面若干张牌，比方说取 n 张牌，然后 [26] 把它们放到这副牌的最底下．然而，在我们的模型中，这最底下可是太远了．因此，让我们变通一下，定义一次 **n – 切牌**（n-cut）是将最上面的 n 张牌同紧接在其下的 n 张牌对换一下．例如，对初始顺序的一次 5 – 切牌将导致

$$6,7,8,9,10,1,2,3,4,5,11,12,13,\cdots$$

如果现在我们从一次 1 – 切牌开始，接着一次 2 – 切牌，依此类推地顺序执行 n – 切牌，那么证明盖伊猜想的正确性将是一件微不足道的事．这是因为 c 号牌在第$[c/2]$次切牌前总是呆在它的老地方，此后每隔一次切牌它就向上移动一步，直至来到最上面，从这儿它又跳了下去，然后再一步一步地回到最上面．因此，为了让事情有趣一些，我们不是单纯的切牌，而是每次切牌

后把最上面的一张扔掉. 于是问题就是,是不是每张牌最终都会被扔掉. 这里的统计行为比起洗牌的情况来有所克制,虽然偶尔也会发一下脾气. 例如,752 号牌坚持了一千九百多万次切牌以后才被扔掉. 然而,看起来这个问题还是可以驾御的. 一张给定牌的轨迹有一个清晰的模式,读者在少数几个例子上花一番工夫即可容易地发现这个模式. 一张牌将经受多少次切牌而不被扔掉的问题,结果归于一个数论中的问题——但这个问题在本文撰写的时候尚未解决.

下面说一点这些问题的来源. 它们都始于金伯林(Clark Kimberling)在《数学难题》(*Crux Mathematicorum*)第 7 卷第 2 期(1991 年 2 月)上提出的一个问题. 金伯林考虑下面的阵列:

$$
\begin{array}{cccccccccccc}
\underline{1} & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & \cdots \\
2 & \underline{3} & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & \cdots \\
4 & 2 & \underline{5} & 6 & 7 & 8 & 9 & 10 & 11 & 12 & \cdots \\
6 & 2 & 7 & \underline{4} & 8 & 9 & 10 & 11 & 12 & 13 & \cdots \\
8 & 7 & 9 & 2 & \underline{10} & 6 & 11 & 12 & 13 & 14 & \cdots \\
6 & 2 & 11 & 9 & 12 & \underline{7} & 18 & 8 & 14 & 15 & \cdots
\end{array}
$$

其中每一行都是用一种蛙跳程序由上面一行而得到的. 从对角线项(已用下划线标出)右边的数字开始,然后是对角线左边的数字,然后回到右边第二个数字,然后左边第二个数字,依此类推,直到你遇上这行的第一个数字. 然后跳回右边,让余下的数字按自然数顺序排列. 一旦一个数字出现在对角线上,就把它排除. 现在金伯林问,"(a) 2 最终会被排除吗? (b) 是不是每个数字最终都会被排除?" 这一过程很容易化成一种洗牌方式,我称之为"金伯林洗牌"(听起来很像 20 世纪 30 年代一种流行舞的名称[①]). 在这种洗牌方式中,当第 n 次洗牌时,先把第 n 张

① "金伯林洗牌"原文是 Kimberling Shuffle,故这里似指 Jitterbug Shuffle,直译是"吉特巴曳步舞",一种在 20 世纪 30 年代风靡美国的社交舞蹈.——译注

牌扔掉,然后把前($n-1$)张牌的顺序颠倒,再一张隔一张地插到紧接其后的($n-1$)张牌中去.

盖伊马上就注意到对(a)的回答为"是".事实上,到第25行(第25次洗牌)时,2就被排除了,用手工计算即可相当容易地予以证明.盖伊还猜想(b)也有肯定的回答.在他那位在剑桥研究计算机科学的孙子安迪·盖伊(Andy Guy)的帮助下,他对1200以下的所有数字验证了这个猜想.列成表格的计算结果表现出与完全的n-洗牌同样的狂野行为.在一次私人通信中,盖伊写道,"我猜想所有的数字终将被排除,但我也猜想没有人打算去证明这一点."因此他实际上作了两个猜想,它们具有如下有趣的性质:如果其中一个被肯定,则另一个就得不到肯定. [27]

一个西班牙语的自描述子

在第2章中萨洛斯开出了一个构造自描述语句的妙方.显然这个过程是与语言无关的,所以这里有一个西班牙语的版本,它是由马德里工业大学计算机科学系的莱马(Miguel A. Lerma)构造的.

ESTA FRASE CONTIENE EXACTAMENTE DOSCIENTAS TREINTA Y CINCO LETRAS: VEINTE A'S, UNA B, DIECISEIS C'S TRECE D'S, TREINTA E'S, DOS F'S, UNA G, UNA H, DIECINUEVE I'S, UNA J, UNA K, DOS L'S, DOS M'S, VEINTIDOS N'S, CATORCE O'S, UNA P, UNA Q, DIEZ R'S, TREINTA Y TRES S'S, DIECINUEVE T'S, DOCE U'S, CINCO V'S, UNA W, DOS X'S, CUATRO Y'S, Y DOS Z'S[①].

我听说萨洛斯也有一个荷兰语的例子.

---

① 这句话译成中文就是:"这个句子正好包括235个字母:20个A,1个B,16个C,13个D,30个E,2个F,1个G,1个H,19个I,1个J,1个K,2个L,2个M,22个N,14个O,1个P,1个Q,10个R,33个S,19个T,12个U,5个V,1个W,2个X,4个Y,以及2个Z."与第2章中的那个英语句子一样,它无法译成具有自描述性质的中文句子.——译注

对一些评论的一个再评论

在爱尔特希(Paul Erdös)众多值得称道的成就中,有一项可能是他在合作论文的数目方面所稳稳保持着的史无前例的世界纪录. 因此,下面这件事会引起人们的兴趣:至少有一篇论文,他是唯一作者,但在某种意义上这又曾经是一篇合作论文. 卡普兰斯基(Irving Kaplansky)刚写完对爱尔希特一篇(合作)论文的评论,就遇上了这位作者本人. 卡普兰斯基说他很欣赏论文的结果,但他不知道主要定理的证明(占了一页半的篇幅)是不是能大大缩短. 爱尔希特作了重新考虑,并且很快就发现这一点确实能做到. 下面摘自《数学评论》(*Math. Review*, Vol. 7, 1946, page 164)的内容提供了这件事的全过程.

Anning, Norman H., and Erdös, Paul. Integral distance. Bull. Amer. Math. Soc. 51, 598—600(1945). [MF12821]

> 作者证明了,对任何 n,在平面上都存在着不共线的点 $P_1, \cdots, P_n$,使得所有的距离 P_iP_j 都为整数;但是不存在具有这种性质的不共线点的无穷集. (参见下一条评论.)I. 卡普兰斯基(伊利诺伊州,芝加哥).

Erdös, Paul. Integral distance. Bull. Amer. Math. Soc. 51, 996(1945). [MF14475]

> 这篇论文全文如下.
>
> "在一篇具有同样标题的注记中(参见上一条评论),证明了平面上不存在一个不共线点的无穷集,其中点与点之间的所有相互距离都为整数.
>
> "可以给下述的推广以一个更简短的证明:如果 A, B, C 是不在一条直线上的三个点,而 $k = \max[(AB, AC)]$,那么最多存在 $4(k+1)^2$ 个点 P,使得 $PA - PB$ 和 $PB - PC$ 都为整数. 这是因为

$|PA-PB|$最大为AB，因此可以假定它的值为0，1，…，k中的一个，这就是说，P在$k+1$条双曲线中的一条上面．同样P也在由B和C所决定的$k+1$条双曲线中的一条上面．这些(不同的)双曲线最多交于$4(k+1)^2$个点．对更高维的情况，一个类似的定理显然成立．"I.卡普兰斯基(伊利诺伊州，芝加哥)．

[28]

因此这里是一个罕见的例子，一篇文章以全文发表在两家不同的杂志上——而现在这样又使它发表在三家杂志上(或许又是一个世界纪录)．

[29]

[30]

第6章 一个有两千年历史的学科的几百个新定理:何处是尽头?

从欧几里得到笛卡儿到 *MATHEMATICA* 再到覆灭?

我们的话题再一次是数学与计算机的相互作用.

计算机能解决数学问题.计算机也能提出问题,而现在看来,它们可能会整个儿地消灭这门学科的一些分支.这就好比当一台计算机成为(如果能够成为的话)象棋世界冠军时,职业象棋就可能遭受灭顶之灾一样.一个潜在的牺牲品就是那个资格最老的数学分支——欧几里得(Euclid)几何,特别是指可以在欧几里得的著作①中找到的几何.

这些想法是由金伯林最近的一些有关三角形各种中心的研究所激发的.中心的例子有形心、内切圆或外接圆的圆心、高线的交点(垂心),等等.这里,"等等"中包括了由金伯林所描述的关于中心的大约91个不同概念.人们可以基本上随心所欲地定义中心,其方法很容易明白.那就是,如果 a, b 和 c 是一个三[31]角形的边长,而 $f(x,y,z)$ 是某种关于 y 和 z 对称的函数,则相应的中心就是这样一个点:它到 a, b 和 c 的距离与 $f(a,b,c)$, $f(b,c,a)$ 和 $f(c,a,b)$ 对应成比例.我们说函数 f 给出了这个

① 这里显然是专指欧几里得为古典几何学作了奠基性贡献的不朽著作——《几何原本》(*Element*).——译注

点的**三线坐标**(trilinear coordinates)[①]. 当然,能引起人们兴趣的是那些对应于自然的几何构造的中心. 在金伯林的清单中,大约有一半的中心曾在文献中出现过,其余则是他的发现(发明). 实验揭示了一件令人惊奇的事:在这 91 个中心中,存在着极其多的共线性. 确实,所有这 91 个点,只要 103 条不同的直线即可覆盖(如果任何三点都不共线,就需要 4 095 条直线).

共线性的经典例子是欧拉(Euler)线,它经过形心、外心和垂心,而且形心位于从垂心到外心的三分之二处. 这个结果作为练习出现在许多初等解析几何的书中. 人们还知道欧拉线经过九点圆的圆心. 如果你想不起九点圆的定义,好吧,那么我也忘了[②]. 但不管怎样,它的圆心位于从外心到垂心的中点. 金伯林发现还有八个中心也在欧拉线上. 应该强调,这些共线性都是这样发现的:用一台计算机对 a, b 和 c 取少数几个数值,并注意到这些不同的中心在准确到小数点后大约第 10 位的条件下排成一直线. 当然,除了一小部分之外,所有这些共线性都是人们前所未闻的,连鬼都不知道. 人们可以把它们看成几百个正在寻找证明的新定理.

举一个例子. **费马点**是一个锐角三角形中与三个顶点的距离之和为最小的那个点. **拿破仑**(Napoleon)**点**(据说是这位皇帝亲自发现的,那时他还没有把兴趣从几何转向征服世界)可这样得到:在一个三角形的每条边上分别向外作一个等边三角形,再将它们的中心分别与原来三角形中与它们相对的顶点连起来,

---

① 粗略地说,一个点的三线坐标就是它与三条给定直线的有向距离,记为 $\{\xi_1, \xi_2, \xi_3\}$. 因此,中心的三线坐标应该是 $\{\alpha f(a,b,c), \alpha f(b,c,a), \alpha f(c,a,b)\}$,其中 α 即比例常数. 由于度量关系往往不是本质的,所以 $\{\xi_1, \xi_2, \xi_3\}$ 与 $\{\beta\xi_1, \beta\xi_2, \beta\xi_3\}$ 可表示同一点,其中 β 是不为零的实数. 还有,三个点的三线坐标所组成的行列式如果值为零,则它们共线. 事实上,三线坐标是射影坐标的一个特例. ——译注

② 三角形三条边的中点、三条高线的垂足和三条高线上从顶点到垂心的中点共圆. 此即九点圆. 可参见本译丛中《近代欧氏几何学》的第 11 章. ——译注

则三条连线共点(定理),这个点就是拿破仑点.

结果发现,拿破仑点和费马点连成的直线不知怎么一来也经过了外心. 这是一个人们可能预见不到的事实. 不过,既然这个现象已被发现,一个证明也就唾手可得. 当然,从原则上说,自从笛卡儿(R. Descartes)发现了坐标几何,这一点就已经肯定,但在实际上,要得到并解出所涉及的多项式方程,可能十分困难,而且会坠入"迷魂阵". 于是像 *Mathematica* 这样的符号运算程序在此登场. *Mathematica* 不在乎"迷魂阵". 给它输入这三个点的三线坐标,在这种情况下它们是

$$费马点:\{\csc(A+\pi/3),\csc(B+\pi/3),\csc(C+\pi/3)\},$$

$$拿破仑点:\{\csc(A+\pi/6),\csc(B+\pi/6),\csc(C+\pi/6\},$$

$$外心:\{\cos A,\cos B,\cos C\}^{①}.$$

其中 A,B,C 是三角形的角. 用 $\sin A\sin B\sin C$ 去乘前两行,以消除分母. *Mathematica* 将这个行列式展开成一个正弦和余弦的多项式,这个多项式结果为零. 证毕.

还有一个奇特之处:存在着一个第二拿破仑点. 作法同上,[32] 只是这次等边三角形作在原来那个三角形的内部而不是外部. 在这种情况下,这个第二拿破仑点和费马点的连线经过的是九点圆圆心而不是外心. 怎样用综合法对此作出证明,我不知道——但 *Mathematica* 对此当然不在话下.

这种状况确实十分奇怪. 什么事情都用计算机来完成. 程序 A 进行了一次探险航行,发现了一大批定理,然后程序 B 接过手来并提供出证明,而当这一切在进行的时候,研究者只要舒服地坐在那儿看着. 这种事应该让机器人来做. 这就使得人们要做一下反思:我们在研究数学时力图成就的是什么? 在一个

① 关于费马点和拿破仑点的三线坐标,其根据可参见本译丛中《近代欧氏几何学》的第 12 章. 至于外心的三线坐标,建议读者自己推导. ——译注

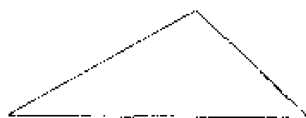
人们工作了两千多年的学科中,突然有几百个新的事实被揭示,这固然不同凡响. 但是数学,或广泛地说,科学,它们所关心的远不是编纂一份关于事实的巨型目录表. 它们希望找到一般的原则,据此能对事实作出推断,而机器人在这方面看来不能有很大的帮助. 它们告诉我们什么是真的,但不能告诉我们为什么. 它们提供大量的信息,但几乎没有悟性.

对三角形的见仁见智

如果你要求人们画一个三角形,他们几乎总是会画出这样的一些三角形:



但几乎从不会画出这样的三角形:



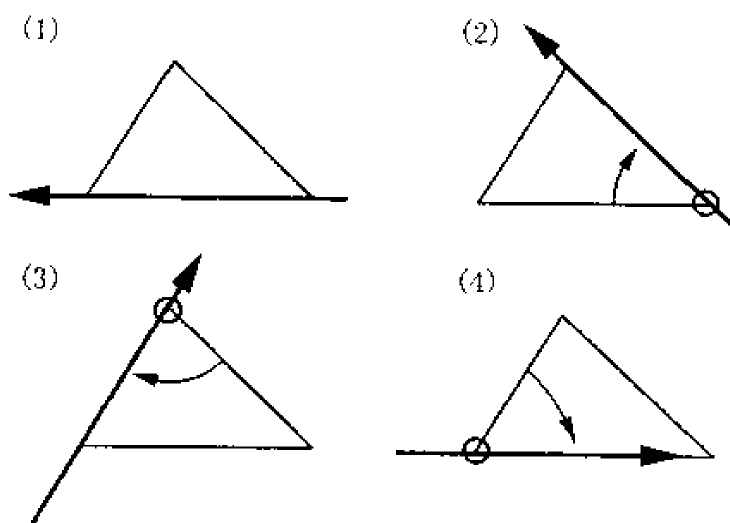
然而,纽曼(Donald Newman)指出,无论你怎么看,一个"随机选取的三角形"是钝角三角形的可能性几乎总是远大于是锐角三角形的可能性. 人们是怎样随机地选取一个三角形的呢? 一种可能的方式是随机地选取三个正数作为三个内角的大小,它们只要满足和为 π 这个条件. 如果这就是选取准则,那么很容易知道,四分之三的三角形是钝角三角形. 这相当于考察 $\mathbf{R}^3$ 中的一个单位 2 - 单形[①],这个单形中坐标都小于 1/2 的点所组成的点集,面积为这个单形的四分之一. 或者,人们可以在单位区间中随机地选择边长,并且问:在那些对应于三角形的三元数组中,对应于锐角三角形的占多大比例? 这里的分析比较复杂,

① 即 $\mathbf{R}^3$ 中由方程 $x+y+z=1, 0<x,y,z<1$ 所决定的三角形. ——译注

但答案结果是 $\pi/4$[①]. 还有一种方式,即随机地在一个单位圆周上选三个点. 这里答案还是四分之三的三角形为钝角三角形. 理由是:这个三角形成为锐角三角形的充要条件是圆心位于这三个点的凸包内部,这相当于把点看成向量时要求圆心是这三 [33] 个点的线性组合,而且所有系数同号. 这种情况四中有一.

三角形与教学

大多数孩子在很小的时候就熟悉了像三角形、正方形和圆这样的几何对象. 怎样利用这种熟悉性,而不是用花去他们如此多时间的算术,来向他们表明一些真正的数学? 这里有几个建议,其中有一些基于我个人的经验. 我想起我第一次遇到一个数学证明时的情形. 那时我五年级,一位同学教我关于三角形内角和的结论,并证明为什么它是对的. 他没有用欧几里得的证明(第1卷,第32号命题[②]). 你会记得,那得先知道平行线为其他直线所截时的情况. 他演示了由下面一列图所示的"移棒式"证明:



① 原文如此,疑有误. 似应为 $2-(\pi/2)$. 还是钝角三角形多一些. ——译注
② 指《几何原本》的卷号和其中的命题编号,下同. ——译注

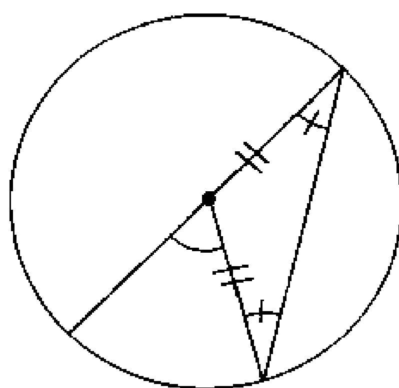
小棒沿三角形底边放置,方向朝西,然后连续转动,先是以右顶点为中心,接着是以上顶点为中心,再后是以左顶点为中心,于是最终它回到底边,但是方向朝东. 因此,用四年级的话来说,它一定是转过了半圈. 我被迷住了——谁又能不被迷住呢? 我天真地想道. 这是我数学经历中曾达到的一个高峰,我的下一个高峰是在几年后我知道了第47号命题,即那个关于斜边平方的定理①的时候. 当然,我那时还没有想到人们可以终生致力于寻找和证明这样的事情(当然,即使不是如此美妙的事情),甚至因做这样的事而得到报酬(有时我仍发现这一点非常重要). 但是这个故事还有下文. 几年前我又思考了这个移棒式论证,因为我醉心于试图把这类在数学中进行的事情传达给外界,包括孩子, 而且我注意到,如果你把那小棒围着三角形接着再转上一次,显然你最后会让小棒的方向再次朝西. 因此当时我问自己,而现在我问你,小棒是不是准确地位于它初始的位置? 它是不是有可能发生了由三角形形状而定的向左或向右的移动? 如果你想对此作一思考,你最好停下来不要读下去,因为我马上就要给出答案.

小棒是位于它初始的位置. 为什么? 你看,第一次围着三角形的周游使小棒的定向发生了反转,而一个反转定向映射总有一个不动点. 但第二次周游与第一次完全一样,它一定也有同一个不动点,因此不会有移动. 接下来,就是今天当我开始继续写这篇文章的时候,我对自己说,好吧,在这根最后回到初始位置的运动的小棒上,有这么一个不动点,那么,它是哪一点呢? 让我们把这做成一个多项选择题. 这个不动点是(a)底边的中点,(b)高线的垂足,(c)底边与对角平分线的交点,(d)除以上三个之外的一个点. 我把答案留给你去想,但附上下述注记:这些文字实在是急就章. [34]

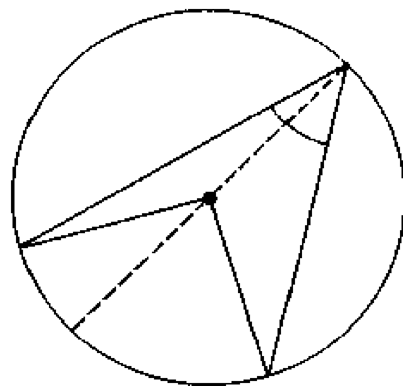
① 即勾股定理,西方称毕达哥拉斯(Pythagoras)定理. ——译注

让我们把话题回到孩子们，这个移棒式的论证显然可以适用于其他多边形．当然，现在不单单是问这根小棒围着这多边形周游一番后是否反了向，还要问它转了几圈．这个问题为卷绕数[①]这个非常基本的概念提供了一个绝妙的直观引导．我们对三角形的不动点论证事实上对所有的奇多边形都有效，但对偶多边形一般无效．

在另一个方向上，一旦你有了第 32 号命题及其推论，即三角形的外角等于它两个不相邻的内角之和，你就可以引入圆，并证明一些结论，比方说，圆周角是圆心角的一半（这还要用到等腰三角形底角相等这个事实（欧几里得，第 5 号命题），但我想大多数孩子会把此作为一个显然的事实予以接受[②]）．如果你忘了这证明的进行过程，下面的图应该唤起你的记忆．



特殊情况

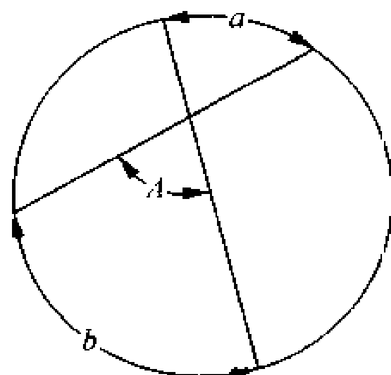


一般情况

① 拓扑学中的一个概念．直观上说，它是指圆周上的一个点在圆周上作了运动后回到起点时其轨迹对这个圆的缠绕圈数．——译注

② 当然，我们大人知道，即使是"显而易见"的命题，也需要有个证明，然而，严格这个概念在五年级肯定没有它的地位．事实上，在任何水平上，需要严格的唯一理由是保证我们不犯错误．严格，就其自身而言，毫无神圣之处．导致"新数学"（20 世纪 60 年代在美国兴起的一场教育改革运动，强调以现代集合论为基础，在严格的逻辑体系下展开中学数学教学．——译注）失败的关键，正是在于没有认识到这个事实．——原注

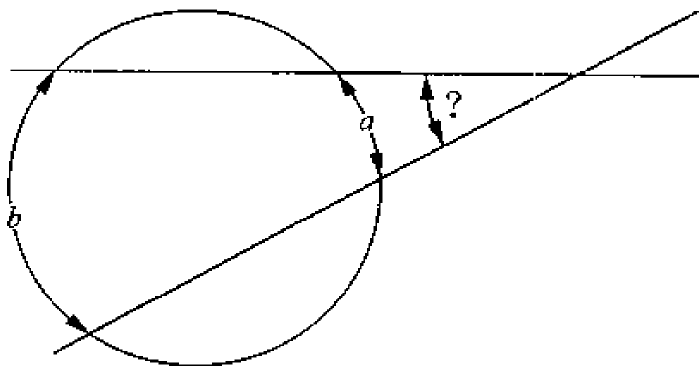
能不能期望普通的 10 岁孩子完成这一切? 或许不能,但我们不试一下是不知道的. 我们怎样才能知道他们是否真的理解了呢? 一个方法是要求他们凭自己的能力做一道类似的题目. 例如,角 A 与弧 a 和 b 的关系是什么?



[35]

(根本没有必要用习惯上的希腊字母去吓唬孩子们. 我记得我自己第一次遇到 ξ 时是如何的惶恐.)这里,作一系列暗示是很有用的. 首先,若他们有困难,就告诉他们答案是什么,并要求他们对此作出证明;如果这不起效,就告诉他们必须再添一条线;再不行,就告诉他们添哪条线,如此等等.

对那些能完成这道题的孩子,下一个挑战可以是这样的:



等等.

现在,美国学童的"数学盲"(一个可怕的表述!)正引起人们

的关注,或许,适量的三角形是一帖良方,在培养比较"有数学能力的"学生的事业中,它会有所益助. 事实上,这个设想的一个最美妙的特点是,它与数完全没有关系.[36]

第7章 协议与大众数学

再说协议:通过电话玩游戏

下面几乎所有的内容由巴拉尼(Imre Barany)提供,在此谨表谢忱.

1. 掷硬币

莫斯科的一位数学家与纽约的一位数学家正在就他们写的一篇合作论文进行远程通话,他们被邀请在巴黎的一个会议上宣读这篇论文.他们两人都想去,但是只有一个人可以去,因此他们决定把这件事付诸运气,这是唯一公平的做法.“好吧,”俄国人说,“我掷了一枚硬币,你说正面朝上还是反面朝上?”“等一下,”美国人说,“这不公平.”“怎么啦?”俄国人说,“你不相信我?那么好吧,我给你一个数,32357650571084661213974696442300978848885300629319084500941003706256554489710395955272354261 69795539,它的最大素因数模4同余于1还是−1,如果你告诉我正确答案,你就可以去宣读论文.”美国人作出猜测后,俄国人揭晓了这个数的素因数分解,它是56123699566021020558766279166381074847903158831451 × 57654165390541998801236990031588314500658098016489,于是美国人可以检验这个乘式的正确性以及因数的素性.

[37]

因此,这是一个可用来代替掷硬币的防欺骗的绝好方法,除非有人发现了因数分解的快速算法(如果真有这种算法的话).

2. 剪刀、石头、布——以及一般的两人对策

上述技巧可以很容易地推广到任何不含随机步骤的两人对策. 例如,假定我们要玩剪刀、石头、布,我就传给你一个大数,它的最大素因数模5或同余于1,或同余于2,或同余于3,其依据是我想出的是剪刀,还是石头,抑或布. 然后你直接宣布你的选择,不必进行编码. 接下来我就把传送过来的那个数的素因数分解揭晓. 这显然对有 n 个策略的对策同样有效,方法是把这些策略用 n 个数分别编码,这 n 个数的最大素因数以一个适当的模分别同余于 $1,2,\cdots,n$. 幸亏有狄利克雷(P. G. L. Dirichlet)[①],我们知道总能找到所需要的素数.

3. 三人或更多人的对策

有趣的是,当有两个以上的局中人时,有关的协议变得极为简单,而且更加安全. 在一个 n 人对策中,每位局中人都有一个给定的有限的**策略**集合,对策中的一**局**是指这些局中人的每一位**同时**从自己的策略集合中挑选出一个元素. 但正是这个同时性的特征,给人们要远程地进行这个对策造成了问题. 解决这个问题的关键在于策略"分享"的机制. 具体地说,如果局中人 i 有 m 个可以选择的策略,那么把它们分别用数码 $0,1,\cdots,m-1$ 编号,并且把这种编号告诉给所有的局中人. 如果现在这位局中人决定采用策略 k,那么他就要选择 $(m-1)$ 个数,并且使这些数的和模 m 同余于 k,再把它们每人一个地传送给其他的局中人. 当每个人都如此做了以后,各人采用的策略就被揭晓了. 显然,如果有谁要想在他所选的是哪一个策略上耍花招,他就会被其他人识破. 并不像两人对策的情况,这个协议是"无条件"安全的. 这意思是说,它不需要假定某些计算在计算机上是不可行的.

---

① 这里指解析数论中著名的狄利克雷定理:设 a,b 是两个互素的正整数,则在等差数列 $ax+b$ 中有无穷多个素数.——译注

4．桥牌

诚然，桥牌以及所有的纸牌游戏，都不属于例3所描述的情况，因为它们不但涉及局中人的策略选择，还涉及大自然的随机运作，即发牌．然而，这里仍有着十分美妙的随机化协议．下面介绍的这个适用于桥牌的协议是由格里戈尔耶夫(Dmitrij Grigoryev)给出的．东家、西家和南家共同商定了对一副牌的一种编码方案，例如用1到52进行的编码．北家并不知道这个编码方案，他给其他三家分别随机地分派了13个数字，于是其他三家就分别知道了自己手中的牌是什么．那些未被分派的数字就是北家手中的牌，当然，他还不知道这些数字代表什么牌．没有关系．要知道他手中是否有黑桃A，他可以抛掷一枚硬币，比方说落下来是正面朝上，他就把这个结果告诉西家．西家如果手中没有这张A，就告诉南家是正面朝上；西家如果有这张A，就告诉南家是反面朝上．一般地说，局中人根据他们手中是否有这张A，或者传递这个消息，或者传递这个消息的"反面"．这样，[38]当且仅当东家告诉北家是正面朝上时，北家就拥有这张A．然后他对黑桃2做同样的事．对整副牌都用这种方式进行了探测之后，北家就得到了他需要的信息，而且什么都没有泄露给其他局中人．

注意到以下这一点是很有趣的：附加一个额外的条件，仍然可以执行上述过程，这个条件就是局中人不允许与他们的同伴直接通话．这样的要求造成了一个问题，因为北家必须给南家发一手牌．这个问题可以如下解决：南家对这副牌创设一套新的编码方案．例如，1是一折，2是一曲，等等．把这套新的编码方案以东家作为传递者传给北家．然后，北家就把这套"一折，一曲，……"中的13个以西家作为传递者发给南家．于是，除了南家，其他人仍没得到什么新的信息．

这个例子提出了一个问题，就是怎样的通信网络可以用来进行这种发牌操作．回答是一个充分而且必要的条件是，这个

网络应是双连通的(doubly connected);也就是说,对每一对局中人,至少有两条不相交的通信线路.人们已看到这个性质在上述例子中是如何被应用的,其中从北家到南家必须用到两条不同的线路.这个条件的必要性意味着,如果这个图有这样一个顶点[①],删去这个顶点就把这个图分为两个部分,那么就不可能进行秘密发牌.我们将在后面回到这一点.

5.扑克

注意这个桥牌的发牌协议依赖于这样一个关键性的事实:整副牌都要发完。对于像扑克那样的一些游戏,这个事实当然不再成立,因此必须设计一个不同的协议.这里是巴拉尼对巴拉尼-富雷迪(Furedi)方法就三个局中人情况的描述.这三个人一个是美国人,一个是俄国人,一个是匈牙利人.美国人有一张俄匈词汇对照表(其中给出了每张牌的名称),俄国人有一张美匈词汇对照表,而匈牙利人会所有这三种语言.现在美国人发给俄国人5张牌,方法是把这些牌的俄语名称和匈牙利语名称传送给他,当然,这些名称对美国人来说是不知所云的.俄国人认识自己拿到的5张牌,并且知道了它们的匈牙利语名称.于是他在他的美匈词汇对照表中设法找出表示这5张牌名称的词,然后把其余牌中的5张牌发回给美国人.最后,他们两人中的任一个都可以给匈牙利人发牌.当然,这仅对一次发牌有效,因为在这次发牌的过程中,美国人和俄国人掌握了一些匈牙利词汇.因此,在每次发牌的时候,匈牙利人必须制造两张新的词汇对照表.

在第4章中,我们看到的不是怎样用协议来玩游戏,而是怎样用协议来不泄露秘密地得到信息.一个典型的例子是一些带

① 这里关于图论的一些基本概念,可参见本译丛中《数学:新的黄金时代》的第7章或其他任何一本图论教材.——译注

有秘密数——例如薪金或者年龄——的人的情况，他们只是对得知这些数的和感兴趣．就像在上面例3中那样，解决这个问题的关键在于分享．每一位局中人都选择$(n-1)$个数，这些数的和分别是他们的薪金，并且把这些数分传给其他人．然后每个人都把他所收到的数的和公布出来，于是这些和的和当然就是原来那些秘密数的和．这个简单的过程还具有一个更重要的特征，即它是“n－保隐私的”．这意思是说，即使这些人的一个真子集合起伙来共享他们的信息，他们仍不能发现关于其他人秘密数的任何情况（除了那些他们知道了和以后总可以得知的信息）．然而，这个和函数是很特殊的，因为在一种恰当的意义 [39] 下，它是唯一可以n－保隐私地计算的函数．因此，举例来说，如果你想知道这些薪金的最高额而不是它们的总和，那么你最多只能做到“少数人保隐私”．这意思是说，只存在这样的一种协议，它使得这些参预者的任何一个人数少于一半的子集都不能通过他们之间的信息共享而得知关于其补集的任何情况．确实，有一个可以被称为协议基本定理的结论，它粗略地断言，任何函数都可以少数人保隐私地计算，但是只有极少数函数，像和函数那样，可以n－保隐私地计算．因此，对于大多数函数来说，如果某个人数达到参预者一半或者超过一半的集团对获取额外的信息感兴趣，那么没有一个协议可以保证隐私不被泄露．我们将就下述无异议问题这种特殊情况来说明这个不可能性结果．要弄懂下面这个由巴拉尼给出的证明，需要有某种较高程度的专注力，但是我想那些愿意在智力上进行这一努力的读者将发现这是值得的．

有一项提交给委员会C讨论的动议，当且仅当任何委员都不投反对票时获得通过．这个委员会只想知道这项动议是否通过，不想知道其他任何信息．我们断言，不管用什么协议，如果这些投票者的集合被划分为两个集团A和B，那么这两个集团中必有一个集团，其成员将获得额外的信息．这表明，举例来

说，如果有7位投票者，那么就根本不可能有4－保隐私的协议．其证明如下．考虑这样一种情况：动议被否决了，在A和B中都有人投了反对票，而且假定这个协议没有泄露任何额外的信息．现在设想有一个局外人，他知道这个协议，也知道动议被否决这一事实，他还得到了所有在A的成员和B的成员之间传递的消息M_{AB}的一个副本．那么必定是这样的情况：这个局外人不可能根据这些信息来排除A的所有成员都同意这个动议的可能性．理由是，如果他能做到这一点，那么B的成员也能做到，因为他们也有条件获得同样的信息．这样他们就能得出A中至少有一名成员投了反对票的结论，而这是不允许他们得知的．这意味着，存在一系列在A的成员之间传递的消息M_{AA}，它们相当于一个当A的所有成员都同意这个动议时按协议在A的成员间所传递的消息的集合．对称地，也存在一个消息集合M_{BB}，它与M_{AB}一起，相当于一个当B的所有成员都同意这个动议时按协议发出和收到的消息的集合．但是M_{BB}与M_{AB}和M_{AA}都相容，因为消息集合M_{AA}不能影响到消息集合M_{BB}的合理性．这样就产生了一个矛盾，因为这表明，当所有的成员都同意这个动议时，会出现一个消息集合，然而它却导致这个动议被否决的结论①．

---

① 这里仅作如下诠释：(1)任何协议都是通过参预者之间合法的消息传递算出最后结果的，每位参预者传出的消息都蕴含着关于他那个秘密数和他所收到的消息的信息，但他们都不可能根据所收到的消息推出其他参预者的秘密数．(2)协议执行到最后，或由全体参预者根据若干个公布于众的消息共同算出最后结果（请参见第4章中那个计算秘密数之和的协议），或由其中一位或几位参预者根据他们所得到的消息算出最后结果并公布于众（请参见第4章开头由那位妇女提出的协议），注意这公布于众的最后结果也是一个在参预者之间传送的消息．(3)不管怎样，在M_{AB}中，总包含着导致最后结果的若干个公布于众的消息，或者最后结果这个消息本身．因此，只要M_{AB}不变，而M_{AB}和M_{AA}及M_{BB}相容，那么它们一起组成的消息集合总导致协议的最后结果．(4)M_{AB}与M_{AA}可能有关联，与M_{BB}也可能有关联，但M_{AA}与M_{BB}没有关联．——译注

把上述推理过程坚持阅读完的读者应该注意到，我们没有在任何地方定义过一个协议甚至一个消息的意思是什么，而且我们只用到这样的一个性质：消息的集合，不管人们认为它的意思可能是什么，总能准确地给出所期望的信息，而不泄露其他任何东西.

最后，让我们回想一下，在秘密地分发一副牌的问题中，我们要求通信网络是双连通的. 这个条件与上面这个不可能性结果具有某种一致性. 具体地说，假定除去某个局中人 P 以后，其余的局中人就分裂为集合 A 和 B，而且他们之间的唯一通信方式是通过 P 传递. 现在，除了那些不在 P 手中的牌之外，其余的牌哪些在集合 A 的局中人手中哪些在集合 B 的局中人手中，这一点必定由在 A 与 B 之间交换的消息所唯一决定. 但是 P 有条件得知所有这些消息，因此，既然他知道发牌的协议，他也就知道了其余的牌是如何分成属于集合 A 的部分和属于集合 B 的部分的，而这是不允许的. 这里的一个推论是，在只有两个局中人的对策中，根本不存在秘密发牌的协议.（另一方面，由阿德勒曼(L. Adleman)、里韦斯特(R. Rivest)和沙米尔(A. Shamir)得 [40] 到的这个领域中最早的结果之一表明，如果人们引进像因数分解那种在计算上困难的问题的存在性，那么两个局中人的情况也可以得到解决.)

参考文献

1. L. Adleman, R. Rivest, and A. Shamir, Mental poker, *The Mathematical Gardner*, Londen: Wadsworth International, 1981[①].
2. Imre Barany and Zoltan Furedi, Mental poker with three or more players, *Information and Control* 59(1983), 84—93.

① 有中译本,《数学加德纳》(戴维 A.克拉纳编,谈祥柏,唐方译,上海教育出版社,1992 年版). ——译注

大众数学

你怎样才能使一个街上的过路人信服“数学的神奇性”呢？有两个颇为吸引人们注意的著名故事，一个是皇帝和麦子的故事，另一个是关于猴子和打字机的故事．第一个故事说到一位皇帝为了向他的顾问表示感谢，允诺予以赏赐，方法是把一些麦粒放在一张国际象棋棋盘的格子里给他，第一天在第一个格子里放一颗麦粒，第 n 天在第 n 个格子里放 2^{n-1} 颗麦粒．问题在于，这些麦子的数量比这个王国的全国麦子年产量还要多（准确地说，是 18 446 744 073 709 551 615 颗麦粒，不过这个数字在最初的传说中可能并没有给出）．我推测，这个故事的寓意是，$2^{64}-1$ 是一个相当大的数，它超出了大多数皇帝或者街上过路人的理解程度。

至于说到那些猴子，通常是六只，它们整天坐在它们的打字机旁不停地“随机”敲打．这里的断言是，“按照概率的法则”，它们迟早会打出莎士比亚的全部剧作，或者，如果你喜欢的话，打出大英博物馆中按字母表顺序排列的所有藏书．经过缜密的推敲，可以发现这是一个相当空洞的断言，因为被如此这般地引用的概率法则，原来恰恰是意味着那些猴子就是做它们被设定要做的事．不过，这种宣称会从听众中赢得一些所期望的“嗬”或“啊”的赞叹声．

在这篇评注中，我将从“数学的奇妙世界”中举出一个实例，它同时涉及到 2 的乘幂和大英博物馆，更妙的是，它是一个准确的、真实的陈述（或者用我们数学家的话来说，是一个定理）．不过，或许在那些没有进入数学之门的人看来，这个陈述十分荒谬．事实上，就是对那些已进入数学之门的人，可能也需要给出一些令人信服的说明．这个陈述是，如果那位皇帝继续把麦子的数量加倍，那么最后总有这样一天，表示麦粒颗数的那个十进制数的开头（最左边的）一大串数码将正好是大英博物馆所有藏

书内容的一个数字编码. 而且, 与那个关于猴子的故事不同, 人们可以十分肯定地确定出一个明确的日期, 在这个日期之前, 那个所期望的数字必将出现. 如果我们把要求稍稍放低一点, 比方说, 问到第几天麦粒的总颗数将以 1992 这些数码开头, 我们可以十分精确地说出这个天数. 它就是 4077, 因为 2^{4077} =
1992013705087952626594071790154694333228996751875655709045
8554902724340590888165604568992842388106505220224508560929
5432660593688618811341717383592195771270360175509747259139
6980975322979705273274284496329986469700577390364451814166
3977330862895710325180204557529520112443619704058048422135
6959873885732461297405047775609547795659185445835292930522
8144517561853986791031728857088119454192862919264872920676
3456064061857571371353807387048488843395986489774707585169
7732623493514078839937944517332643907216940073010180734522
2345381190866138147712882678330900614392266934143305874301
9071586844344804050363054126636829830293768181717125429760
6540696641659184042791404285088026651195083237967528435722
3361275519123621702432220487417524596137967663790511344054
7833845268843347369847734151454764458652171217413606969170
0605213431340405977821755020789277607513227165339662058438
4298171659566062245134648926394234907237821419023374001203
2631697281797408521686200442571129300122590123365398330613
9144796219210553134483041738024837350879350484938262831338
0338301788753216880314585298318696631470332063679582455858
0605967490865295637699213466298243312409364759942739543939
6364873350091299670503904558487495571467978179185913811628
0251318272.

[41]

为了下文的引用, 我们指出, 上面这个数字中数码的分布是十分均匀的. 其中只有 118 个 0, 有 137 个 3, 而平均的个数是

124. 但是这种偏离均匀性的误差完全是在通常允许的范围之内.(然而,有一件事情我还不能从数学上作出解释,那就是4077正好也是我社会保障号码的最后四位数.)

用比较正式的术语,这个定理是说,对于任何一串十进制数码 N,总存在一个自然数 n,使得 2^n 的开头一串数码就是数码串 N. 这个结果甚至可能有商业上的潜在价值. 把你的出生日期、月份和年份组成一串数码寄给我们(随信寄上 d 美元),你将在回信中收到你个人的 n,保证与你相伴终身.

这个定理的另一美妙特征是,它的证明只需要你知道怎样做长除法,再加上一个连一位皇帝也应该能理解的数学事实,那就是与狄利克雷的名字相联系的著名的**抽屉原理**. 这个原理说,如果你把一亿亿(或者你能想像得到的巨大数目)零一个物体放进一亿亿个盒子里,那么必定有一个盒子中放了多于一个的物体.(皇帝:你的意思是你非得到数学的研究生院去学**那东西**吗?)这个原理是怎样在目前这种情况中得到应用的? 好吧,我们从使得 2^n 至少包含 $d+1$ 个数码的第一个 n[①]开始. 于是,在接下来的 $10^{d+1}+1$ 个2的乘幂中,你一定会在什么地方找到两个不同的数,它们的开头 $d+1$ 位数码完全相同. 现在把其中较大的数除以较小的数,并回忆一下长除法的规则. 你得到的数将有两种情况:(a)首位数是1,接着是至少 $d+1$ 个0[②](再后面是其他东西);(b)开头是 $d+1$ 个9[③].(你遇到的情况是(a)还是(b),取决于那两个数的第一个不相同的数码谁大谁小.)让我们把在情况(a)中得到的商记为 Z,我们宣称,如果连续取

---

① 作者使用的符号似有混淆,这里的 n 并不是我们要证明其存在的 n. 顺便代作者补充说明一点:假设数码串 N 有 d 个数码. ——译注

② 原文如此,似有误. 只能保证有 d 个0. 例如设 $d=1$,容易找到 $2^5=32$ 和 $2^{15}=32768$,它们开头的两位数码相同,但 $2^{15}\div 2^5=1024$,1后面只有一个0. 好在这个差错并不影响后面的结论. ——译注

③ 同样,只能保证有 d 个9. ——译注

Z 的乘幂，那么这些乘幂展开为十进制数后，它们开头 d 位数码组成的数字将按自然顺序依次把所有的 d 位数都展现出来．这是因为，如果把任何数 N[①]乘以 Z，那么所得的积要么是开头 d 位数码与 N 的完全相同，要么只是第 d 位数码比 N 的大 1．要明白这一点，可考虑这样一个数 Z'，它的位数与 Z 相同，首位数是 1，接下来是 d 个 0[②]，接着又是一个 1，其余都是 0．现在，Z' 大于 Z，而且它显然具有上面楷体字所述的性质：因为如果 $N = ab\cdots xy\cdots$，其中 x 是第 d 位数，那么 $N\times Z'$ 的计算由

$$\begin{array}{l} a\,b\cdots x\,y\cdots 0\cdots 0 \\ \underline{+\qquad\quad a\,b\cdots} \\ a\,b\cdots \end{array}$$

42

给出，而积的第 d 位数将依据 $y + a$ 是大于 9 还是不大于 9 而为 $x+1$ 或 x．对情况(b)的论证类似，只不过这里所涉及的是借位而不是进位．

注意这个证明是构造性的，它给出了一个从 N 找出 n 的算法，但是这证明给出的上界是一个天文数字．对于一个 d 位数，这上界大约是 2^{10^d}．事实上，理论和实践都表明，如果我们寻找的是第一个有这种性质的 n，那么可以发现，平均说来 n 大约与 N 有同样的量级．我们为那些对此颇有兴致的读者提供一个 *Mathematica* 程序，它是海尔梅尔(Stephan Heilmayr)写的，可用来找出这个最小的 n：

```
fract[x_]: = If[x = = Floor[x],1,x - Floor[x]];
power2[x_]: = Block[{Z = fract[Log[10,x]],n = 1,
u = fract[Log[10,x + 1]],pow = Log[10,2],t = pow},
While[N[t < Z] || N[t> = u],pow = pow + Log[10,2];
```

① 同样，这个 N 并不是我们的那个数码串 N．——译注

② 原文如此，似有误．应是 $d-1$ 个 0．——译注

t = fract[pow]; n = n + 1]; n][1]

n 事实上较小，这是理所当然的，因为我们处理的是一种一致分布现象．更准确地说，$\log_2 10$[2]的倍数对模1是一致分布[3]．利用这个事实，我们可以推导出关于2的乘幂的一些更进一步的性质．例如，在其十进制展开中含有一百万个连续的7(或者含有任何其他给定序列)的2的乘幂，在由全体2的乘幂所组成的集合中的密度[4]为1[5]．我感谢费尔德曼(Jack Feldman)给我指

---

① 其基本思想是：定义函数 fract(x)($x>0$)如下：当 x 是整数时为1；当 x 不是整数时为 x 的小数部分．然后对输入的 x 按序逐个检验 lg2 的倍数，直到找出这样一个倍数 $n$lg2，使 fract(lg($x+1$)) > fract($n$lg2) ≥ fract(lg$x$)，于是 x 作为数码串将出现在 2^n 的十进制表示的开头，而且这个 n 是最小的．但这里似有一个缺陷：当输入的 x 是10的乘幂时，由于 fract(lg$x$) = 1，任何的 n 都不会使 fract($n$lg2) ≥ fract(lg$x$)．其实这时 n 只要满足 fract(lg($x+1$)) > fract(n lg2)即可．——译注

② 原文如此，似有误．根据上述程序，应为 lg2．——译注

③ 所谓一个数列对模1是一致分布，是指对任何$(\alpha,\beta)\subset(0,1)$，设该数列的前 N 项中其小数部分落入(α,β)的项的个数为 $f(N)$，则有 $\lim_{N\to\infty}(f(N)/N)=\beta-\alpha$．详情可参见《数论导引》(华罗庚著，科学出版社，1975年版)．粗略而平均地说，若一个数列对模1是一致分布，则这个数列的小数部分将在(0,1)中各处"均匀地"出现．请读者据此及上述程序体会 n 较小这个事实．——译注

④ 设 A 是自然数集的一个子集，$B\subset A$，$A(x)$，$B(x)$分别是 A，B 中不大于正数 x 的元素的个数，则 B 在 A 中的密度是 $\lim_{x\to\infty}(B(x)/A(x))$．可参见《数论教程》(J.-P.塞尔著，冯克勤译，丁石孙校，上海科学技术出版社，1980年版)．——译注

⑤ 这里有个证明思路．不妨考虑含有一个7的情况．先看7首先出现在第一位(左数，下同)的2的乘幂，它们十进对数的小数部分将落在(lg7，lg8)中．lg2的倍数对模1是一致分布，故它们在全体2的乘幂中的密度为 lg8 − lg7．再看7首先出现在第二位的2的乘幂，它们十进对数的小数部分落在(lg1.7，lg1.8)，…，(lg6.7，lg6.8)，(lg8.7，lg8.8)，(lg9.7，lg9.8)之一中(注意其中没有(lg7.7，lg7.8))．这些区间两两不交(与(lg7，lg8)也不交)．于是相应的密度为这些区间长度之和．再看7首先出现在第三位的2的乘幂……依此类推，可得一系列两两不交的区间，含有一个7的2的乘幂在全体2的乘幂中的密度即为所有这些区间长度之和．把(0，1)上的勒贝格(Lebesgue)测度由单调增函数 $f(x)=10^x$ 转换为(1，10)上的勒贝格－斯蒂尔切斯(Stieltjes)测度，即可证明这个长度之和为1．这里关于测度的概念，可参见《集合与面积》(李惠玲等著，广西教育出版社，1999年版)．——译注

出这一点. 对于"几乎所有的2的乘幂,即除了有限个外,都含有一百万个连续的7,或至少含有一个7"这种更强的说法,情况如何呢? 也就是说,在其十进制展开中一个7也没有的2的乘幂是不是有无穷多个? 如果你认为这不可能,那么考虑这样的事实:"差不多所有的"非负整数都含有一百万个连续的7. 这意思是说,具有这种性质的非负整数的集合的密度为1. 其理由如下. 注意在所有的一百万位数中,不是全由7组成的数所占的比例为$(1-10^{-6})$[①],于是在所有的n百万位数中,第一段一百万位数码不全是7,第二段一百万位数码不全是7,……第$(n-1)$段[②]一百万位数码也不全是7的数所占的比例为$(1-10^{-6})^n$[③]——你明白了吗? 在另一方面,当然存在着无穷多个非负整数,它们一个7也不含有. 换句话说,统计上的信息与这类问题实在是没有关系. 作为另一个实例,回想一下在2^{4077}这个数中,各个数码大约以十分之一的密度出现,这种行为很可能是典型的,由此人们可能倾向于得出这样的结论:或许从某处开始,所有的2的乘幂都至少含有一个7. 但是,如果我们考虑按自然顺序排列的全体非负整数的序列,那么根据大数律,那些其中所含的7的百分数小于,比方说小于9.99的数的密度又为零[④]. 然而,却又存在着大量的不含7的非负整数.

---

① 似应为$(1-10^{-10^6})$,而且是在不多于一百万位的数中占的比例. ——译注

② 似应为第n段. ——译注

③ 似应为$(1-10^{-10^6})^n$,而且是在不多于n百万位的数中占的比例. ——译注

④ 其证明大意如下:对任取的一个非负整数,记其右数第i位上的7的个数为X_i,则X_i是一个独立同分布的随机变量序列,它们都以概率0.9取1,以概率0.1取1,且$EX_i=0.1$. 设A_n为事件"一个不大于10^n的非负整数所含的7的百分数小于9.99",B_n为事件"$(X_1+X_2+\cdots+X_n)/n<9.99\%$",则$A_n\subset B_n$. 故$P(A_n)\leqslant P(B_n)$. 但是$P(A_n)$就是在不大于$10^n$的非负整数中所含7的百分数小于9.99的数所占的比例,而$P(B_n)<P(|((X_1+X_2+\cdots+X_n)/n)-0.1|>0.0001)$. 据辛欣大数律,当$n$趋于无穷大时,后者趋于零. ——译注

在无论如何都不能从我们通常的公理得出证明的意义下，是否存在着无穷多个不含 7 的 2 的乘幂这个问题或许是"不可回答的". 然而，它应该是真还是假呢？当然，这依赖于人们的数学哲学信念. 或者，作个异想天开的假设，会不会有人打算证明这个问题等价于某个不可思议的基数的存在性？到了这一步，这个问题就变成了一个神学上的空谈而不是数学问题了——看来我们最好就此打住.

[43]

一个非数学问题

1. 说出美国唯一的不具有印第安名称的中西部州的州名（中西部的意思是，在阿巴拉契亚山脉和落基山脉之间，但不包括墨西哥湾沿岸的南部各州）①.

[44]

2. 为什么在一个关于数学的专栏中会有这个问题？

① 对于不熟悉美国地理和历史的读者，这里不妨给出译者的答案. 这个州似应是 Indiana——印第安纳州. Indiana 系新拉丁语，意思是"印第安人的土地". 这样你就不难答出下面第二个问题了.——译注

第8章 变分方法的六种变分

思　　想

许多年前，作为密歇根大学的一名物理学研究生，我十分幸运地听了斯廷罗德（Norman Steenod）的函数论课程．这从根本上改变了我的人生道路．我听了那门课程以后，再加上几次私人谈话，我决定从物理学转向数学．我特别记得有一次斯廷罗德在给我们上课时，为了试图描述数学研究是怎么回事，他说在整个这门学科中，事实上只有大约十几个思想被人们所一再反复地使用，而一旦你掌握了这些思想，你就可以说是圈中人了．我后悔当时没有在惊讶之中保持镇定，请他给出他那张思想表中的前十二个思想．无论怎样，我想在任何人的思想表中，有一个思想应该是人们所称的变分方法[①]．

变分方法

变分方法通常用来给出存在性证明．为了试图证明一个具有某种性质的对象的存在性，人们挑选出一个使某个函数达到极大值或极小值的对象，于是所得到的对象就可被证明为具有 [45]

① 这里所说的变分方法（variational method）与求解泛函极值问题的变分法（calculus of variations，一译变分学）不是一回事，虽然前者有时也涉及泛函的最大值或最小值，如下文提到的使狄利克雷积分取最小值的函数．——译注

所期望的性质．证明方法是，如果这个对象不具有那种性质，人们就可以对它进行“变更”，使得那给出的函数继续增大或继续减小．这样的例子在数学中处处稠密[①]．最令人熟悉的或许是罗尔(Rolle)定理，而最具有历史意义的则可能是黎曼对著名的映射定理的证明[②]．这是他通过使狄利克雷积分达到最小值而予以证明的．我们知道，黎曼的证明有一个严重的缺陷，他没有证明这个最小值的存在性．直到几年之后，才由希尔伯特(D. Hilbert)成功地补出了这段遗漏的论证．

当然，这个方法可以追溯到更早的年月．最初等的例子或许是算术基本定理的标准证明，这要追溯到欧几里得．其中关键的一步是证明任何两个正整数 a 和 b 都有一个最大公因数，也就是说，有一个 a 和 b 的公因数 d，它能被 a 和 b 的所有其他公因数所整除．

最大公因数

由于是在找某种最大的东西，人们会预料这将涉及到一个最大值问题的解决．结果并非如此，正确的途径不是解决一个最大值问题，而是一个最小值问题．具体地说，就是寻找使得 $d = ma + nb$ 的最小正整数 d，其中 m 和 n 都是整数(这里关于最小值的存在性是没有问题的，事实上，它等价于一条佩亚诺(Peano)公理[③])．a 和 b 的任何公因数都整除 d 这件事于是立即

---

① 一般拓扑学中的术语，这里是一种比喻用法，表示无处不在．——译注

② 黎曼映射定理可以如下粗略地表述：对于复平面上任一个不是全平面的单连通域 D，存在一个在某种意义下唯一的共形且一一的映射 f，使得 $f(D) = \{z: |z| < 1\}$(开单位圆盘)．几乎任何一本较完全的单复变函数教材都会提到这个定理．但黎曼当年的证明现在却难以寻觅．——译注

③ 这是指关于自然数的佩亚诺公理系统中的归纳公理．可以证明，归纳公理等价于如下的“最小数原则”：自然数集的任何一个非空子集都有一个最小的元素．有兴趣的读者可参见《数学归纳法》(华罗庚著，上海教育出版社，1963 年版)．——译注

得证,但是还得证明 d 本身也是 a 和 b 的一个公因数,于是"变分"在此登场.不妨假设 d 不整除 a,则根据欧几里得算法,存在一个小于 d 的正整数 r,使得 $a = qd + r$.于是 r 将是 a 和 b 的一个更小的整系数线性组合,这就产生了一个矛盾.

这个证明有一个显著的特征,它像许多变分方法的证明一样,立即导致了一个关于找出两个数的最大公因数(gcd)的简单算法.

我对这个人们非常熟悉的证明作了十分细致的考察,我想指出,它不仅仅是变分方法的一个应用.它是一种**对偶定理**的一个实例.我们看到,整除 a 和 b 的**最大**正整数同时也是一个能表示成 a 和 b 的整系数线性组合的**最小**正整数.这类定理尤其在线性规划理论中占据着中心的地位.我们将在后面回到这一点.

西尔维斯特问题

变分方法的另一个特征是,它经常导出非常简短的证明.这方面的一个惊人例子是西尔维斯特(J. J. Sylvester)于 1893 年提出的著名问题:设 S 是平面上的一个有限点集,且任何经过其中两个点的直线都一定经过其中另一个点,证明这些点都在一条直线上.不论是西尔维斯特还是他的同时代人都没能找到一个证明.过了将近 50 年,才由加莱(Gallai)发表了第一个证明,但相当复杂.下面这个简短的证明现已广为人知,它是由凯利(L. M. Kelly)于 1948 年发现的(见《美国数学月刊》(*Amer. Math. Monthly*) 55, p.28).假设这些具有西尔维斯特所述性质的点不共线.每条经过其中两点的直线 L 和不在这条直线上的一个点 p 都组成一个线点对(L, p),在所有这些线点对中选取一个使得从 p 到 L 的距离 d 为最小的.令 q 为从 p 向 L 所引垂线的垂足.于是(变分)根据假设,在 L 上至少存在三个点 a, b 和 c.因此其中两个点,比方说 a 和 b,将以 a, b, q 的顺序位于 q [46]

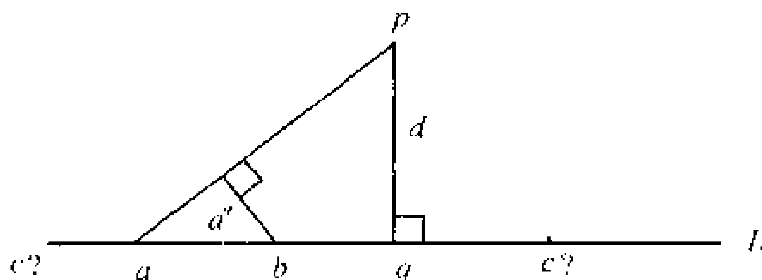


图 8.1 凯利的证明

的同一侧(c 可在任何一侧),如图 8.1. 但这样从 b 到直线 ap 的距离 d' 就小于 d,这就产生了一个矛盾.

凯利的证明确实简短——但是,这里有考克斯特(H.S.M. Coxeter)的说法(见《几何学导引》(*Introduction to Geometry*), Wiley, 1961, p.181):“这件关于共线性的事[西尔维斯特的问题]显然属于序几何. [确实,在复数域或有限域上这个结果不成立!你可以轻易地在环面上发现一个九点的反例.]凯利的欧氏几何式证明涉及外在的距离概念;这好比用一把长柄锤子去砸一个杏仁. 真正恰当的坚果钳由下述证明所提供.”

考克斯特的可爱的证明(很高兴这个证明用的也是变分方法)依赖于帕施(Pasch)公理. 这条公理以它最简单的形式断言:一条直线不可能只与一个三角形的一条边相遇.(理解这条公理的一个方法是,把它看作若尔当(Jordan)曲线定理[①]的一个非常初等的特例. 如果这条直线通过一条边进入这个三角形,那么它必定要穿过另一条边以回到外面来.)这个证明的图与凯利证明的图非常相似,不过这次我们选取的是任意的点 p,并找出一条从它出发的射线 R,R 上没有 S 中其他点,但至少与一条连接 S 中点的直线相交. 每条这样的直线都与 R 相交于某点,

① 若尔当曲线定理是说:平面上的任何简单闭曲线都把这个平面分成两个区域. 详情可参见《拓扑学奇趣》(伏·巴尔佳斯基,伏·叶弗来莫维奇著,裘光明译,湖南教育出版社,1999 年版).——译注

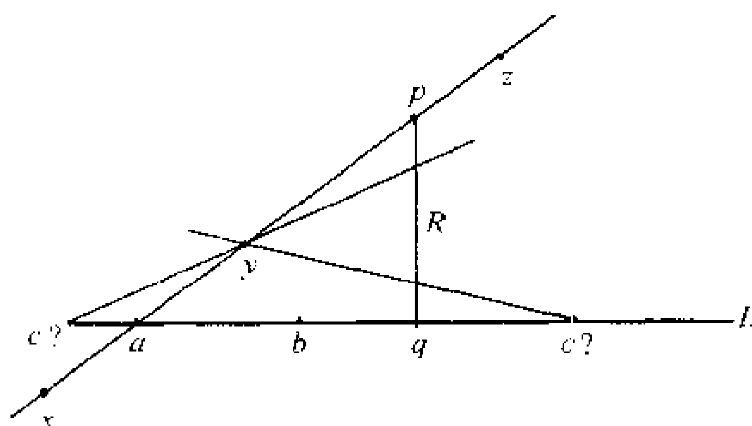


图 8.2 运用帕施公理的证明

于是我们可从中选取一条直线 L,它与 R 的交点 q 距 p 最近(当然,不是在距离的意义下,而是把它们看作 R 上的一个序集. 也就是说,在 p 和 q 之间没有其他交点). 现在(变分)L 上一定有两个点 a 和 b,它们位于 q 的同一侧. 我们证明直线 ap 上不可能有 S 中的另一个点. 有两种情况.

情况 Ⅰ 另一个点 y 位于 a 和 p 之间. 于是(如图 8.2)不论 c 在哪里,将帕施公理应用于三角形 apq,直线 cy 将与 R 相交于一个比 q 更近于 p 的点.

[47]

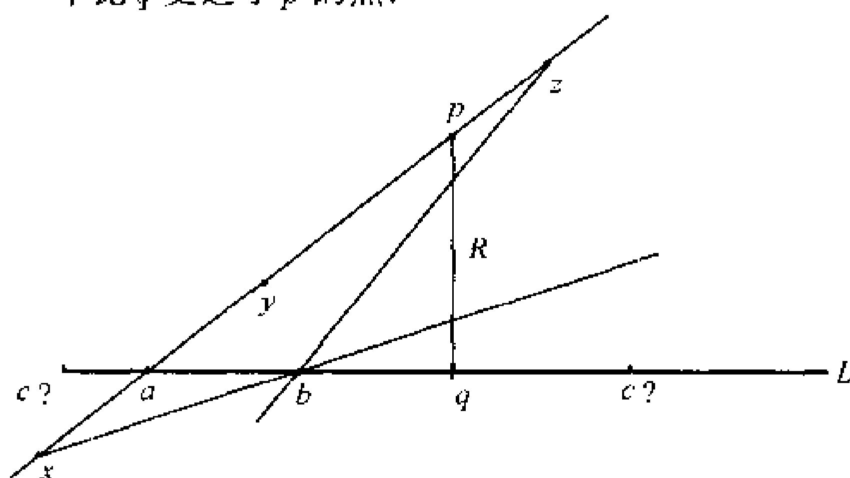


图 8.3 运用帕施公理对另一种情况的证明

情况 Ⅱ　另一个点 x 或 z 不在 a 和 p 之间. 于是,同前面一样,不是 bx 就是 bz,将与 R 相交于一个比 q 更近于 p 的点.(见图 8.3.)Voilà[①]!

然　　而

平面上有呈一般分布的 $2n$ 个点,其中 n 个是红色的,n 个是蓝色的. 试证明它们能连成 $2n$ 条**不相交的**线段,每条线段一个端点是红色的,另一个是蓝色的.

当然,现在你已掌握个中要义,知道你应该用这样一种方式来选取线段:使得这些线段的长度之和为最小. 于是这些线段将不会相交,因为(变分)如果它们中有相交的,就用欧几里得《几何原本》的第 20 号命题(表明一个红点和一个蓝点之间的最短距离是沿着一条直线的那个命题)来证明,你可以让它们变得不相交,从而使它们的总长度减小. 然而我们又在用距离这把"长柄锤子"来对付一个显然关于仿射(但非射影)不变量的定理. 不过我想希腊人恐怕并未对此感到过困惑.

伯克霍夫的台球

令 T 是一个具有光滑(C^1[②])边界的凸图形台球桌. 那么对于任何 $n>1$,都存在一条具有 n 个反弹点的周期性台球轨迹(G·D·伯克霍夫(G.D.Birkhoff)的定理).

证明　在所有的内接 n 边折线形(当然,它们的边可能相交)中,选取一个其周长达到最大值的. 这将是一条台球轨迹. 也就是说,在每个反弹点上,入射角将等于反射角,因为如果不是这样,就将反弹点沿边界朝着同切线形成较大角的边稍稍移动一下. 这将使周长增大(一道美妙的微积分练习题,可放在教

① 法语,意为"就是这些".——译注

② C^1 表示具有连续的一阶导函数.——译注

科书上通常称为"相关变率"的部分[①]).■

注意这个结果对于偶数边折线形是平凡的,因为台球只是沿着这集合的一条直径[②]来回反弹.

承蒙洛托准许,图8.4给出了一张椭圆形台球桌上的两条具有5个反弹点的伯克霍夫台球轨迹(是用 *Mathematica* 绘出 [48]

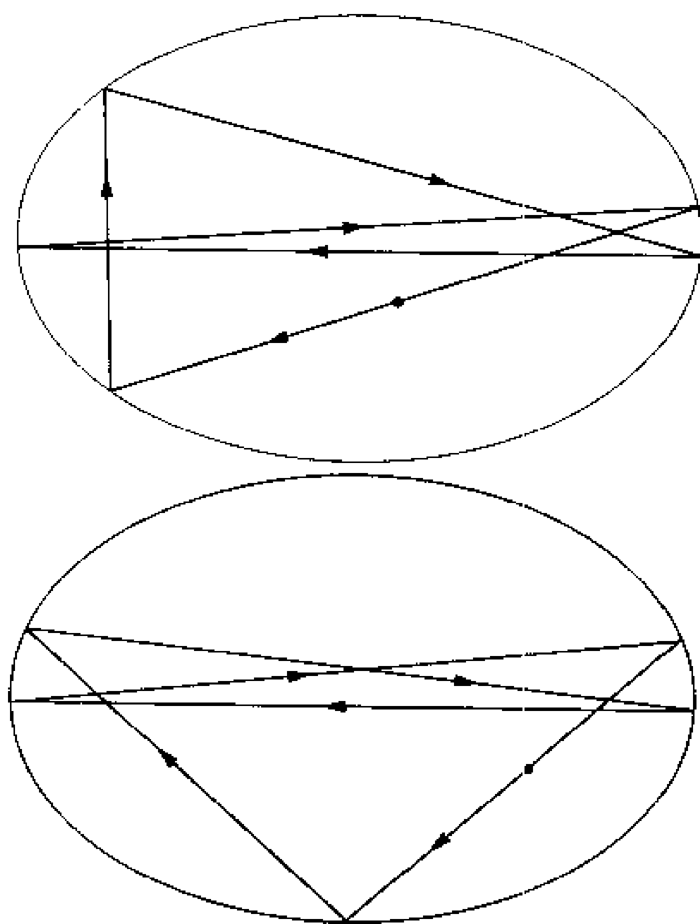


图8.4 一张椭圆形台球桌上的两条伯克霍夫台球轨迹

---

① 设在某反弹点 r 处,边 a, b 同切线形成的夹角分别为 α, β,,则可证当 r 沿边界朝着 b 的方向移动一个 ds 时,$a+b$ 的增量为 $ds(\cos\alpha-\cos\beta)$.——译注

② 对于平面上的一个有界凸图形来说,直径就是连接其边界上相距最远的两点的线段.——译注

的). 有趣的是,关于非光滑边界的台球桌,特别是多边形台球桌,人们知道的并不多. 例如,对于锐角三角形,总存在着一条具有3个反弹点的轨迹,反弹点就在各条高的垂足处,这次是通过寻找具有最小周长的内接三角形来解决的.(你能证明吗?)已经证明,对于各个角都是π的有理数倍数的多边形,存在着周期性的轨迹. 然而,关于各个角是π的无理数倍数的钝角三角形,显然什么都不知道.

废除种族隔离定理

说到简短的证明,这里有一个纽曼给出的例子,它是由克拉姆金(Murry Klamkin)讲给我听的.

> 给出任何一个图,试证明总可以对其每个顶点着上白色或黑色,使得每个白色(黑色)顶点的邻点中至少有一半是黑色(白色)的.

我必须承认,我当时不知道如何证明它,但是普罗普(Jim Propp)证了出来,而且他的证明不到两行,只用了六个单词. 这[49]个证明在本章的末尾给出.

稳定指派定理

最后是一个来自经济学的例子. 设有n名工人和n名雇主,如果工人i为雇主j工作,他们将合作生产出a_{ij}个单位的某种货物,比方说面包. 假定每位雇主只能雇用一名工人,那么谁将为谁工作,而且劳资合作双方将怎样把他们生产的面包划分为工人的**工资**w和雇主的**利润**p?(工资和利润是以面包的形式而不是以金钱的形式给付,是为了强调我们是在处理具有其真正价值的货物,而不是纸币和硬币.)有一个简单的经济学上的"均衡"条件给出了回答. 如果在一种指派方式中,某位工人i挣得了工资w_i,而某位没有雇用i的雇主j赢得了利润p_j,那么

我们就需要有

$$w_i + p_j \geqslant a_{ij}, \tag{1}$$

因为如果这不等式不成立，那么 i 和 j 就会实行合作，从而使他们两人都可能获得更多的面包．于是问题就成为：是否总存在一种将工人指派给雇主的方式和一种对每对劳资合作双方所生产的面包进行划分的方式，使得(1)被满足？这样的一种配置方式称为一个稳定指派或**均衡指派**．

在考察这个存在性问题之前，我们请大家注意均衡指派的一个奇妙的性质，它是有时人们所称的“经济学基本定理”的一个特例．显然，对于整个社会的福利来说，劳资指派关系应该使得他们生产出的面包达到尽可能最大的产量．这样的一种指派称为**最优的**．

定理 均衡指派是最优的．

证明 为了记号上的方便，我们假设在一个均衡指派中，工人 i 就为雇主 i 工作，因此有

$$a_{ii} = w_i + p_i. \tag{2}$$

对(2)就 i 求和，得

$$\sum a_{ii} = \sum w_i + \sum p_i. \tag{3}$$

现在考虑任何一个其他指派，其中工人 i 被指派为雇主$\sigma(i)$工作．根据均衡条件(1)，

$$\sum a_{i\sigma(i)} \leqslant \sum w_i + \sum p_{\sigma(i)} = \sum w_i + \sum p_i. \tag{4}$$

最后一个等式成立是因为 σ 是一个一一映射．根据(3)和(4)，就得到 $\sum a_{ii} \geqslant \sum a_{i\sigma(i)}$，这正是我们要证的最优性．（正是因为有着像上面这样的定理，经济学家到处在颂扬“市场经济”的优越性．）■

于是我们看到，这个均衡性质同一个非常自然的最大值问题相关，这提示应该用这个最大值问题来证明均衡指派的存在

性. 真是愚人节开的玩笑! 结果恰恰相反,就像在最大公因数的情况那样,人们应该去考虑"对偶的"最小值问题. 在所有满足稳定性条件(1)[①]的 $2n$ 元组$(w_1,\cdots,w_n,p_1,\cdots,p_n)$中,选取一个使得 $\sum w_i+\sum p_i$ 达到**最小值**的. 就这些使前式达到最小的

[50] w 和 p,考虑由所有使得(1)作为等式

$$w_i + p_j = a_{ij} \tag{5}$$

而被满足的序偶(i,j)所组成的集合 S. 我们定义,如果(5)对一个序偶(i,j)成立,则称 i 和 j 是**相容的**. 如果现在可以从这个集合中挑选出一个由 n 个不相交的[②]相容序偶组成的子集,那么容易知道,这就给出了所期望的均衡指派.

现在,这个证明中的关键一步用到了霍尔(Philip Hall)那著名的"配对定理"[③]. 如果不存在这样的子集,那么就一定有一个"瓶颈". 具体地说,就是有一个由 k 个雇主组成的集合,与这些雇主相容的工人的总人数将小于 k. 但是在这种情况下(变分),根据"供求规律",人们将对这些"供不应求"的工人的工资

---

① 应当指出,条件(1)在引进时,是仅针对没有实行合作的工人 i 和雇主 j 而言的(对已实行合作的工人 i 和雇主 j, 则肯定有 $w_i+p_j=a_{ij}$);而这里所要满足的条件(1),是针对所有涉及指派的工人和雇主而言的(谁为谁工作还不知道). 因此略有不同.——译注

② 所谓两个序偶(i,j)和(i',j')不相交,是指 $i\neq i'$ 和 $j\neq j'$ 同时成立. 事实上,这相当于一个2分图中的两条边不相邻. 可参见《图和网络及其应用》(费培之编著,四川大学出版社,1996年版).——译注

③ 关于组合数学中的配对定理,涉及"相异代表组"的概念. 设 $A_1,A_2,\cdots,A_n$ 是非空集合 S 的 n 个子集(其中可以有相同的子集,可把它们称为一个"集系"),若对 $i=1,2,\cdots,n$,存在 $x_i\in A_i$,x_i 各不相同,则称元素组 $x_1,x_2,\cdots,x_n$ 为集系 $A_1,A_2,\cdots,A_n$ 的一个相异代表组. 配对定理是说,集系 $A_1,A_2,\cdots,A_n$ 有相异代表组的充要条件是,对任何的正整数 k, $1\leqslant k\leqslant n$,以及任何的下标组 $i_1,i_2,\cdots,i_k$, $1\leqslant i_1<i_2<\cdots<i_k\leqslant n$, $A_{i_1}\cup A_{i_2}\cup\cdots\cup A_{i_k}$ 至少有 k 个元素. 详请可参见《组合数学》(H.J.赖瑟著,李乔译,科学出版社,1983年版). 在这里,可设与雇主 j 相容的工人所组成的子集为 A_j,便可得到文中所说的结论.——译注

w_i 增加 ε，这将使得那 k 个雇主的利润 p_j 有所减小，从而使得 $\sum w_i + \sum p_i$ 减小，这就产生了一个矛盾.

就像在我们的其他例子中那样，这个证明导致了一个找出一个均衡指派的好算法（$O(n^2)$[①]），即所谓匈牙利算法，它是由库恩（Harold Kuhn）给出的.

这里是普罗普对废除种族隔离定理的证明：

Maximize the number of interracial neighbors[②].

请注意又一次由证明导致了一个好的着色算法. 它最多用 E 次迭代即可完成计算，这里 E 是这个图的边数. 具体地说，从任意一个着色方案出发，如果有一个顶点不符合条件，就改变它的颜色.

（这篇文章写好后，吉文撒尔（Alexandre Giventhal）提出了一个三个词的证明：Maximize ferromagnetic energy[③].）

---

① 粗略地说，$O(n^2)$ 表示这个算法的运算量在 n^2 的量级. n 是工人或雇主的人数. 在计算复杂性理论中，凡一个算法的运算量在所算问题的规模（如这里的 n）的多项式（如这里的 n^2）量级上，就称这个算法是好算法.——译注

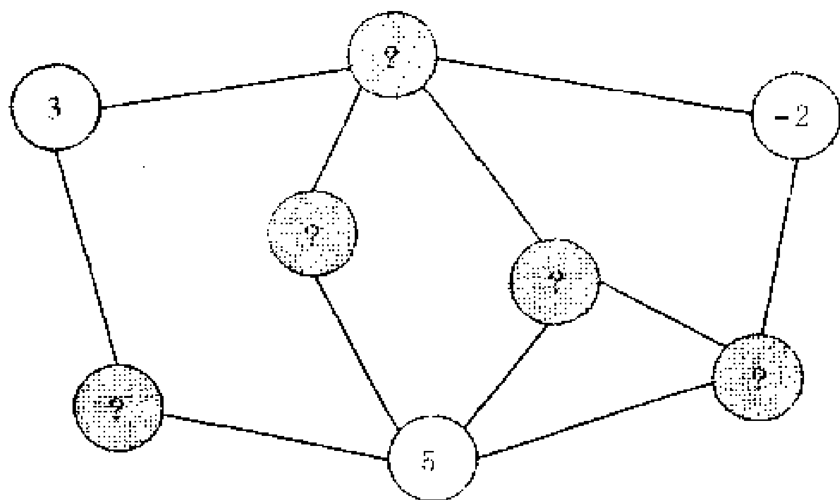
② 其中文意思是："使不同种族的邻居数达到最大值."这里白色顶点和黑色顶点被比作不同种族的白人和黑人，图中相邻的两个顶点被比作邻居. 由此也可知为什么称之为"废除种族隔离定理".——译注

③ 其中文意思是："使铁磁能量达到最大." 这里看来是把一个图看成是一个多原子系统，每个顶点代表一个原子，邻接的顶点看作近邻的原子，每个原子的未满壳层是最外层轨道，且这个轨道上只有一个电子. 根据海森伯的铁磁性理论，该系统的能量由单电子能量、库仑能量和铁磁能量（近邻原子的电子自旋间的交换作用能之和）组成，而铁磁能量为 $-\frac{2A}{\hbar^2}\sum_{i>j(\text{近邻})}\boldsymbol{\sigma}_i\cdot\boldsymbol{\sigma}_j$. 其中 A 为近邻原子间的交换积分，对于铁磁性系统，$A>0$；$\hbar$ 为普朗克常量；$\boldsymbol{\sigma}_i$ 和 $\boldsymbol{\sigma}_j$ 分别为第 i 个原子和第 j 个原子上电子的自旋角动量；求和仅限于近邻原子的电子对. 显然，当电子自旋角动量方向相反的近邻原子对最多时，该系统的铁磁能量最大. 详情可参见《铁磁性理论》（姜寿亭编著，科学出版社，1993 年版）.——译注

关于变分方法的补遗

当然,可用变分方法解决的问题说也说不完. 这儿有一个讨人喜欢的合二而一的例子,它是由克利福德·加德纳推荐给我的. 这个例子有一个特别的长处:它提供了微积分的一个简单而美妙的应用,可以安插在传统的大学一年级课程的第三周左右进行讲授.

请考虑**离散热流问题**. 给出一个图,如下图所示. 其中有一些结点保持在固定的温度上(3,5 和 -2). 关于热传导的定律要求其余的每一个结点都符合**平均值性质**. 具体地说,就是每个余下的结点的(稳态)温度应该是它所连接的结点的温度的平均值.



[51]

问题 这样的一种温度集合是否总是存在? 如果存在,它是否唯一?

(当然,这是一个解线性方程组的问题,但是我们的学生要到大学二年级才能学到这个内容.)

解答 存在性:令 t_i 是结点 i 上的温度. 考虑函数 $f(t_1,\cdots,$

$t_n) = \sum(t_i - t_j)^2$,求和范围是所有邻接的结点对(这就是这个系统的热能). 选取使这个函数达到最小值的各个 t. 要明白这些值满足所期望的条件,只要注意到,如果除了 t_k 以外,所有结点的温度都固定在这些使这个函数达到最小值的值上,那么 t_k 必定使作为单变量函数的f 达到最小值,令关于 t_k 的导数为零,即得所证.

啊?你说,但我们怎么知道这个最小值一定存在? 我的回答是,数学在其两千年的发展过程中,从未为此类问题担心过,因此没有理由把它们强加给大学一年级学生. 对于那些想主修数学的学生,当他们到大学三年级时,有的是时间来迫使他们为这些事情操心,虽然在有些情况下他们会怨气冲天.

唯一性(用最大值原理!):假定存在两个符合平均值性质的温度集合,于是它们的差也将符合这个性质,而且在那些保持固定温度的结点上温度之差为零. 现在考虑一个温度之差达到最大的结点,于是根据平均值性质,它所有的邻接结点一定也在这个温度上. 同样,它所有邻接结点的邻接结点也是如此. 这样下去,我们最终将到达一个保持固定温度的结点(我们假设这个图是连通的),因此这个最大值为零. 对于最小值的情况同样如此.

Q.E.D.

经我花言巧语,反复劝诱,伯克利数学科学研究所行将退休的所长同意将下面这篇杰作奉献出来,这或许反映了他在任职期间所遇到的一些事情.

造诣颇深的年轻人

《帕欣丝》中邦索恩所唱歌的一个模仿之作,在此向吉尔伯特爵士(Sir William Gilbert)和沙利文爵士(Sir Arthur Sullivan)表

示敬意及歉意[①].

卡普兰斯基作

在现代数学的世界中,
如果你想开山辟道,
你得口若悬河,无所不晓.

你必须显示你懂得
拓扑和艾达尔上同调,
物理学中的弦,也要提到.

偏微分方程的一个集合
无非是一个闭路空间,
因此你可以活得,愉快悠闲.

[52] 你应做得略显活泼随和,
如果他们把量子群苦参,
就告诉他们,这很平凡.

合唱:
你告诉他们你的妙方[②],
于是人人喜气洋洋.
因为你使他们看到
根据布尔巴基
只要人们努力如常,
那么双曲流形和仿射平面
就是小波化了装.

---

①《帕欣丝》(*Patience*),又名《邦索恩的新娘》(*Bunthorne's bride*),著名讽刺性轻喜歌剧,由英国剧作家吉尔伯特和作曲家沙利文于1881年所作.邦索恩和帕欣丝为剧中男女主人公名.——译注

② 原文是scheme,一般情况下指"规划"、"方案"等,但在代数几何学中又用作术语"概型".此处是双关.——译注

当然你将就庞加莱猜想
向大伙儿作个演讲;
满嘴含糊,用词夸张.

但不要暗示你能拿下,
有人干过而后悔终生,
大量彩色粉笔,化作飞尘.

你要为纽结引进
不变量多多,
让引理汹涌而来,激浪高波.

如果他们变得睡眼惺忪,
就发明一个精妙的理论,
使费马大定理,不再烦人.

合唱(吉尔伯特的原词):
你说话举止酷似神魔,
于是人人会说:
如果这位年轻人表达见解
用的术语在我看来不知说什么,
嗬!这位造诣颇深的年轻人
一定是造诣深得谁也没见过. [53]

[54]

第9章 铺砌环面 切蛋糕

本章我们将专讲关于游戏数学中两个极著名问题的一些最新结果，这两个问题颇有点历史．第一个问题用不同的正方形铺砌曲面，第二个是为公平地分蛋糕而设计程序．关于铺砌的结果相当完整，而关于切蛋糕的工作尚处于比较初级的阶段．

用不同的正方形铺砌曲面

这个问题是，或更确切地说，本来是：哪一种矩形能被大小不同的正方形所铺砌？图 9.1 例示了一种被 9 个这样的正方形所铺砌的 32×33 矩形．这个例子看来是莫伦(Moron)于 1925 年发现的，它出现在鲍尔(Ball)的《数学游戏》(*Mathematical Recreations*)[1]和施坦豪斯(Steinhaus)的《数学万花镜》(*Mathematical Snapshots*)[2]中．1940 年，塔特(Tutte)、布鲁克斯(Brooks)、史密斯(Smith)和斯通(Stone)证明这是这种例子中"最小的"(见《杜克数学杂志》(*Duke Math. J.*) 7(1940)，312—340)．这意思是说，没有一种矩形能被少于 9 个的正方形如此铺砌．但他们还证明，另外恰有一种矩形也能被 9 个正方形铺砌，即图 9.2 所示的 [55] 61×69矩形．这些作者反复搜寻，终于发现了一种能被如此铺

① 有中译本，即本译丛中的《数学游戏与欣赏》．——译注

② 有中译本，《数学万花镜》(胡·施坦豪斯著，裘光明译，湖南教育出版社，2000 年版)．——译注

砌的正方形．要找一篇引人入胜的普及性文章，可见马丁·加德纳(Martin Gardner)的《关于趣题和娱乐题的第二本〈科学美国人〉图书》(*Second Scientific American Book of Puzzles and Diversions*, Simon and Schuster, 1961)中由塔特写的“化方为方”(Squaring the Square)那一章．

现在，矩形的这种“化方”可以用通常所用的把对边视作同一的方法平凡地转化为柱面、环面、默比乌斯(Möbius)带或克莱因(Keine)瓶①的化方．但是这些曲面也有着非平凡的化方，甚至有着**单化方**(simple squaring)，这就是说，在这种化方中，铺砖的任何一个子集的并都不是一个矩形．然而，不久以前，还不知道这些曲面是否会有所需正方形少于 9 个的化方．后来，在 1991 年，布雷斯韦尔(Bracewell)发现了默比乌斯带的一种只用了 8 个正方形的化方．最近，这个问题为查普曼(S. J. Chapman)彻底解决．或许最简单但最惊人的结果是，一条 1×5 的默比乌斯带能被 2 个正方形所铺砌，这从图 9.3 可看得十分清楚．要看懂这个例子，人们必须推广铺砌的概念，即允许那个把正方形映入曲面的映射在正方形的边界上重叠．

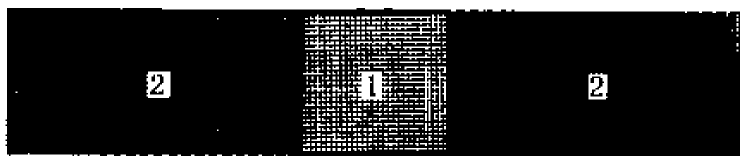


图 9.3

① 这些都是拓扑学中的典型曲面，其中默比乌斯带和克莱因瓶的直观图象，可参见本译丛中《数学趣闻集锦(上)》的 45—47 页．但读者应当习惯一种用矩形的表示方法．若把一个矩形的一对对边同方向地视作同一(即把这矩形卷起来使对边粘合)，就是一个柱面；若把一个矩形的一对对边逆方向地视作同一(即把这矩形的一条边扭转 180 度再与其对边粘合)，就是一条默比乌斯带．对于用矩形如此表示的一个柱面，若再把另一对对边同方向地视作同一，就是一个环面；若再把这另一对对边逆方向地视作同一，就是一个克莱因瓶．当然也可以把用矩形如此表示的默比乌斯带的另一对对边逆方向地视作同一，这就是所谓“射影平面”．——译注

迄今考虑的都是带有一条边界的曲面,这迫使人们把那些正方形放置得有一条边同那边界平行. 环面和克莱因瓶的情况就不再是这样了. 如果正方形可任意方向地放置,则事实上任何两个正方形都可铺砌某个环面. 具体地说,令两个正方形的边长分别为 a 和 b,考虑把一个边长为 c 的正方形的一对对边视作同一而得来的环面,其中 $c^2 = a^2 + b^2$. 图 9.6 表明了这种铺砌的方式,同时提供了毕达哥拉斯定理的一个新(?)证明.

图 9.7 给出了一种更为对称的形式.

如果只允许铺砖平行于这个正方形的边,则已证明环面没有非平凡的 9 - 化方. 默比乌斯带的任何一种化方都给出了克莱因瓶的一种化方. 如果只用 6 块或 6 块以下的铺砖,这些就是克莱因瓶的仅有的化方,但如果用 7 块或 8 块铺砖,情况就不知道了. 如果铺砖不一定要平行于那个大正方形的边,那么克莱因瓶的铺砌是否存在,对此人们也是不甚了了.

查普曼解决这些问题的技巧与塔特等人所用的十分不同,也比他们的简单. 他主要是用元素仅为 0,1 和 - 1 的矩阵对铺砌作了一种巧妙的编码.

分　蛋　糕

有一只蛋糕,要分给我们这 n 个人. 我们的口味各不相同. 有些人喜欢上面的糖衣,另一些人则偏爱那巧克力馅心,如此等等. 是不是存在一种给我们每人分一块蛋糕的方法,使得每个人都认为自己得到了一块在各人所得中最为理想的蛋糕(这样的一种分配被称为**无妒忌的**(envy-free))? 唔,那得看情况. 首先,那只蛋糕的表面必须不能过于崎岖不平. 如果我们所有人都觊觎那颗镶在中间的樱桃,那么除非能用某种方法把这颗樱桃切开分给我们,不然是无计可施的. 这个关键的性质,即任何分下来的蛋糕块都能划分成更小的块,对应于这样一种思想:我[58] 们的口味可用所谓"无原子的"测度来表示,具有可列可加性,等

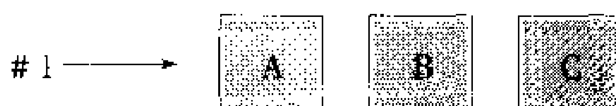
等,等等. 在这个模型中,如果分下来的一块蛋糕是指任意的可测子集,那么有着一个非常强的存在性结论. 不但存在着一种无妒忌的分配,而且存在着一种我们都认为包括自己在内的每个人恰好得到了那蛋糕的 n 分之一的分配. 用另一种方式叙述,就是存在着一种把那蛋糕切成 n 块的方法,使得我们都对自己得到哪一块毫不在乎. 它们看上去同样美味可口. 这个由杜宾斯和斯帕尼尔(Spanier)于 1961 年证明的事实,是由李雅普诺夫(Lyapunov)的一个中等强度的著名定理导出的. 这个定理是说,一个向量测度的值域是凸的.

然而,作为一件需要实际操作的事情,蛋糕的任意可测块是不可能如此容易地得到的. 于是,斯特罗姆奎斯特(Stromquist)于 1980 年创建了一个比较现实的模型. 在这个模型中,可以把那蛋糕理解为一个区间,而分下来的蛋糕块则被规定为子区间. 或许一种长方体的大面包是对这种情况的一个比较适当的图解. 利用一种不动点定理,可以证明无妒忌的分配总是存在的. 可是,众所周知,不动点定理是非构造性的,因此,对人们如何才能把这美味的食品瓜分得人人都心满意足,斯特罗姆奎斯特的结果没有给出任何启迪.

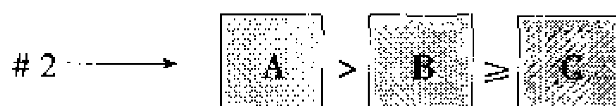
对这个问题有着一个截然不同的处理方法,就是不仅要求证明无妒忌分配的存在性,而且要求给出一个导致这种分配的程序,或者说一个**协议**(下面我们将如此称呼). 一个可作为原型样本的例子是关于平凡的两人情况的程序:我们中的一人把这蛋糕分为两个部分,另一人则从中挑选出他喜爱的那个部分. 这种方法具有显然令人十分满意的性质,如果我们中有谁最终认为自己吃了亏,那么他只能责怪自己. 长期以来,设法为 n 人的情况设计具有这种性质的协议,一直是一个悬而未决的问题. 沿着这种思路的最好结果是属于塞尔弗里奇(John Selfridge)的一个精妙的三人协议. 下面我们即予描述. 我们把局中人记为 #1, #2 和 #3.

三人协议

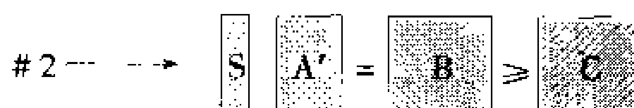
第一步：#1把蛋糕“三分天下”，分成对他来说都可以接受的3个部分.



如果#2和#3所喜爱的蛋糕块不相同，我们就完事了. 如果不是这样，比方说他们都喜爱A，而#2不但喜爱A甚于喜爱B，而且认为后者至少与C一样讨人喜欢. 于是，



第二步：#2从A上削下一个“长片”(S)，剩下A′，使得A′和B对他来说都可以接受：



第三步：#3在A′，B和C中挑选出他所喜爱的一块.

情况Ⅰ　#3挑选了A′. 那么#2就挑选B，而#3挑选C(至此没人产生妒忌).

[60]

余下来的事是分那个长条. 现在即使#3把整个长条都拿去，#1也不会产生妒忌. 除了这一点，这件事就同我们原来的问题一样.

第四步：#2三分S.

第五步：#3挑选，然后#1挑选，再后是#2.

情况 Ⅱ　#3 没有挑选 A′．那么 #2 就把 A′拿下．接下来的程序中，除了这次是由 #3 三分蛋糕，由 #2 首先挑选外，其余就同前面的程序一样．

这个程序有许多美妙的性质．首先，它很节约，最多只需要切 5 次．其次，就像那个“我切你选”的协议，它对局中人所能做的事作了最少的假设．具体地说，(1) 给出任何一个蛋糕块和一个正整数 k，假设一位局中人能把这个蛋糕块分成 k 个对他来说都可以接受的子蛋糕块；(2) 如果一位局中人喜爱一块蛋糕甚于喜爱另一块蛋糕，那么他能从前一块蛋糕上削下一部分，使得剩下的部分和另一块蛋糕对他来说都可以接受．最后，喜爱次序只要满足**弱可加性**，也就是说，如果 A，B，X 和 Y 是不相交的蛋糕块，而一位局中人喜爱 A 甚于喜爱 X，喜爱 B 甚于喜爱 Y，那么他喜爱 $A \cup B$ 甚于喜爱 $X \cup Y$．

迄今为止，即使是对四位局中人的情况，还没有发现一个能满足所期望条件的协议．然而，最近布拉姆斯(Steven Brams) 和泰勒(Alan Taylor) 的工作表明人们正在取得进展．这两位作者给出了他们所谓的一个有限算法，据此可以得到一个无妒忌的分配[①]．不过，他们的程序十分复杂，而且看来要求局中人能够精确地测定出任何蛋糕块的价值．况且，即使是四位局中人的情况，对于可能需要的切蛋糕次数也没有一个事先估计的上界．作为一个例子，如果对某位局中人来说蛋糕块 A 的值比蛋糕块 B 大了一百万分之一，那么用他们所建议的算法就可能需要切一百万次才能达到所期望的分配．人们也许希望对一般情况能设计出简单得可同塞尔弗里奇的协议相比较的程序．另一方面，已经有人猜想这样的协议并不存在．一个悬而未决的有趣问题．

① 关于布拉姆斯和泰勒的结果，有兴趣的读者可参见译者在《科学》1996 年第 1 期上的介绍性文章《从所罗门王的智慧谈起》及其所引文献．——译注

分馅饼:一个未解决的问题

一个分配纵然是无妒忌的,但可能有着另一些不太令人满意的性质.举一个例子,假定你我打算分一个长方体的大面包,我们还是把它比作一个区间,而且假定在我们两人中无论谁的测度下,这个面包关于其中点是对称的.那么,如果我们在这个中点把它一分为二,我们就得到了一个斯帕尼尔-杜宾斯分配.在这个分配中,我们俩都认为我们每人正好得到了那面包的一半.然而,假定我生性喜欢硬面包皮,因此我特别想得到那面包的两头,而你却最好不要这些部分.于是,如果我们以某种方式把这面包一分为三,你拿中间,我拿两头,那么我们俩的所得就都更为理想了.

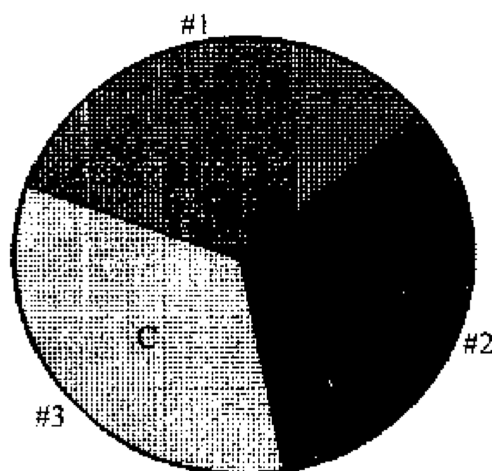
一般地,我们说一个分配是**次要的**(dominated),是指存在另一个分配,它使所有局中人分到比他们在前一分配中的所得更为喜爱的部分.显然,最终的分配如果既是非次要的又是无妒忌的,那将是十分令人满意的.于是有一个一般的问题:在一个给定的模型中,这两个条件是否总能同时得到满足?在这里,要请读者回忆一下斯特罗姆奎斯特的模型,在这个模型中,所有分下来的部分都被规定为一个区间的子区间(这不同于前面那个例子,在那里我所分到的是两个不相交区间的并).事实上,对于这种斯特罗姆奎斯特分配,我们有下面的定理. [61]

定理 一个无妒忌的斯特罗姆奎斯特分配自动地是非次要的.

证明 令 P 是一个把那区间分成 n 个子区间的无妒忌划分,而 Q 是另外任何一个也把那区间分成 n 个子区间的划分.于是,如果 P 和 Q 是一个区间的不同划分,那么必定存在 P 下的某个区间 I,它严格地包含了 Q 下的某个区间 J(用一分钟对此想一下).但是这样一来,在分配 Q 下得到 J 的不管是谁,他的

所得都不会比他在 P 下的所得更为理想,因为他喜爱 I 的程度至少同喜爱 J 一样,而他喜爱他在 P 下所得的程度至少同喜爱 I 一样. 后者的理由是 P 是无妒忌的.■

这个定理把我们引向了馅饼问题. 假定有一个馅饼,要分给三个人,而且规定分下来的馅饼要具有传统的形状,也就是说,是扇形.



是否必定总存在着一种分配,它既是无妒忌的又是非次要的?

我们都会犯错误Ⅱ

作为对我征求著名数学家犯错误例子的响应,有几位读者提到了勒贝格的一个众所周知的错误. 下面这段精彩的描述是由林德(Doug Lind)给出的.

"勒贝格曾试图证明平面上一个博雷尔(Borel)集[①]的投影是直线上的一个博雷尔子集[见《关于可解析表示的函数》(Sur le fonctions représentable analytiquememnt),载《数学杂志》(*Journal de Mathématiques*),Series 6,Vol.1(1905),p.195]. 他的论证用到

① 简单地说,博雷尔集就是在 n 维欧氏空间中可从开集出发,用至多可数次取余、并、交的运算而得到的点集.——译注

了这样一个‘事实’:如果在平面上有一个单调减少的集序列[1],那么它们的交的投影就是它们的投影的交(在以‘Supposons que e soit F de class 1 ou 2’[2]开头的第三段文字中). 这当然是错误的(没有一个一年级研究生会犯这样的错误!),但它却导致产生了关于解析集[3]的苏斯林(Suslin) - 鲁津(Lusin)理论. [暂停 20 秒,请读者找出一个明显的反例.]

"鲁津甚至要求勒贝格为他关于解析集的著作[即《解析集及其应用教程》(*Leçons sur les Ensembles Analytiques et leurs Applications*)]写了一篇序言. 其中勒贝格说道,这是他所犯的最富有成果的错误!" [62]

① 即满足条件 $A_1 \supset A_2 \supset \cdots \supset A_i \supset A_{i+1} \supset \cdots$ 的一列集合 $A_n(n = 1, 2, \cdots, i, i+1, \cdots)$. ——译注

② 法语,意为"假设 e 为第 1 类或第 2 类的 F". 这里 F 可能是 fonction(函数)的词头缩写. ——译注

③ 设 E 为 n 维欧氏空间,则 $E \times [0,1]$ 中的博雷尔集在 $[0,1]$ 上的投影就是解析集. 但解析集的概念要广得多,这里仅是一个特例. ——译注

第10章 自动机蚂蚁 不用圆规的作图

勤劳的蚂蚁

本书的大多数读者无疑对康韦那著名的生命游戏①十分熟悉．这个游戏中的"生命"是所谓"胞腔自动机"的一个例子．我在这里将讨论另一个这样的自动机，叫做"蚂蚁"．虽然它描述起来十分容易，但它的行为却十分有趣，而且有点儿神秘．

这个蚂蚁生活在被坐标网格划分成一个个"胞腔"的平面上．胞腔有两种，一种是白色的，一种是黑色的(后面我们还将引进灰色的胞腔)．最初，这个蚂蚁呆在一个称为原点的胞腔内，头朝着东南西北四个方向之一．它开始按下述规则一个胞腔一个胞腔地旅行：它爬向它朝着的那个胞腔．当它爬进的是一个白色(黑色)胞腔时，它就向右(向左)转 90°，而这个胞腔也由白变黑(由黑变白)．规则就是这些．给定了黑白胞腔的某种分布后，游戏就这么开始了，且看这个蚂蚁如何动作．

设胞腔最初都是白色的，蚂蚁则不妨头朝东．这个特例是[63]一般情况的一个典型．在蚂蚁旅行的初期，大约开头 500 步内

① 一张坐标方格图上分布着一些"生命"，一个小方格内至多一个．其生灭规则是：若一个生命有 2 或 3 个生命做邻居，则它可存活，否则灭亡；若一个空白的小方格周围恰有 3 个生命，则一个生命在此小方格中诞生．这就是生命游戏．其意义在于，如此简单的规则竟产生了千变万化的生灭过程．读者可到 http://psoup.math.wisc.edu/Life32.html 去下载一个叫做 Life32 的软件来观赏这个游戏．——译注

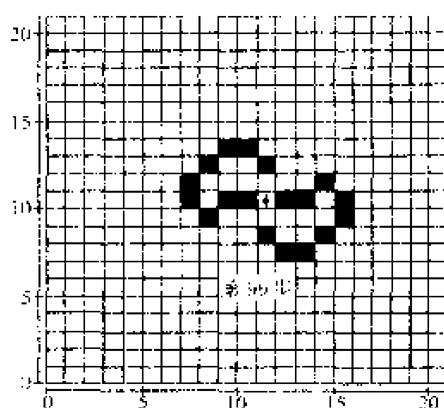


图 10.1

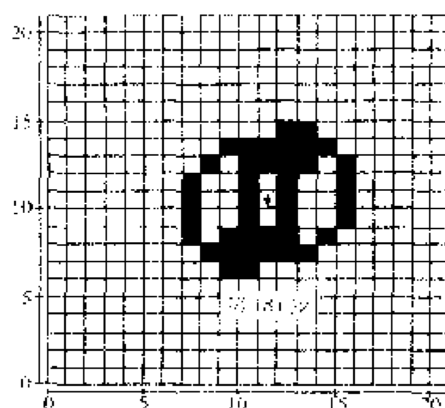


图 10.2

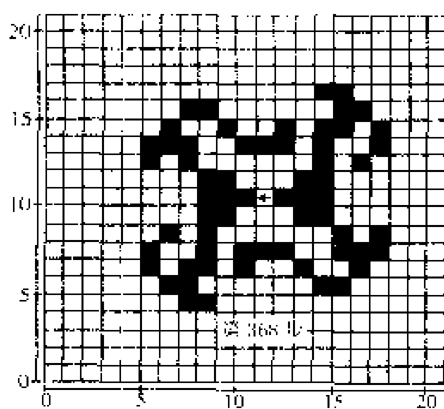


图 10.3

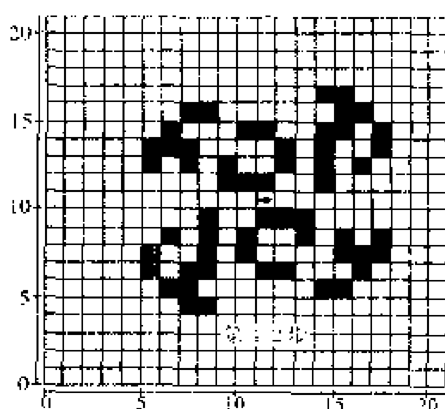


图 10.4

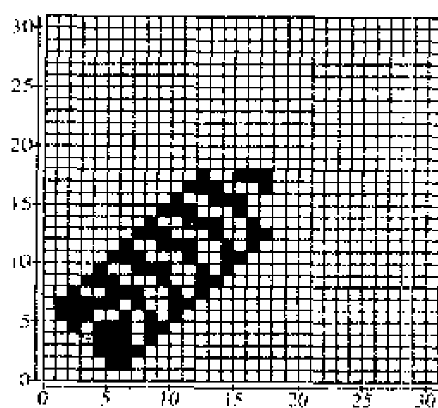


图 10.5

[64]

吧,它不时地回到原点,留下由黑白胞腔组成的中心对称图案,如图10.1~图10.4所示.据我所知,还没有人对为什么会出现这些图案作出一个解释.然而,过了一会儿,局面就变得相当紊乱,这种状况大约持续了10000步左右.但此后这个蚂蚁似乎突然决定了它要去的地方,于是义无反顾地朝着正西南而去,留下了如图10.5所示的那个周期性图案.最早要人注意这个现象的普罗普把这个图案称为一条公路.上了公路,蚂蚁走过104步,就来到它刚才出发点西南两个单位的地方,然后便无休止地重复这个过程.

如果最初有一些黑色的胞腔,那么这个蚂蚁当然会循着另外一条路线行进,但是在几百次试验中,它最后总是沿着四个可能的对角线方向之一建造出一条公路.这种情况是必定会发生的吗?没有人知道.但是有一个十分迷人的结果,它是由布尼莫维奇(L. A. Bunimovich)和特罗别茨科伊(S. E. Troubetzkoy)证明的.

定理 一个蚂蚁的轨道总是无界的.

证明 首先注意到这个蚂蚁的旅行规律是可逆的.黑白胞腔的图案和蚂蚁的当前位置及朝向决定了它来自哪个胞腔.而如果一条轨道是有界的,那么总有一个黑白图案迟早会重复出现,因此,根据上述注意到的事实,这条道路一定是周期性的,从而每个被访问过的胞腔一定会被无穷多次地访问.现在,请注意一个关键的事实,就是这个蚂蚁的爬行是沿水平方向和沿垂直方向交替着进行的.这意味着这个平面上的胞腔可以划分成h胞腔和v胞腔,就像国际象棋棋盘那种式样,前者总是沿水平方向被访问(即从左边或者从右边进入),后者总是沿垂直方向被访问(即从上边或者从下边进入).现在考虑一个被这个蚂蚁访问过的"最大的"胞腔M,这意思是说,在这个胞腔的上方和右方没有一个胞腔被这个蚂蚁访问过.假定M是一个h胞腔,

那么根据它的最大性，蚂蚁一定是从左边进入而从下边出去的，因此这个胞腔原来一定是白色的．但是它接下来就变成了黑色，于是下一次蚂蚁从左边进入时，就必须从上边出去．这同最大性发生了矛盾．如果 M 是一个 v 胞腔，则证明过程类似．

很巧妙，你说呢?

这个游戏的一种变化是引进第三种类型的胞腔，一种灰色的胞腔，它们的性质是：当爬进这样一种胞腔时，蚂蚁只是沿它原来的方向继续行进．灰色胞腔不改变自己的类型，它们永远保持着灰色．注意在这个模型中，上述无界性定理的证明失效了，因为再也不可能把胞腔划分为 h 胞腔和 v 胞腔．而且科恩(E.C.D.Cohen)确实发现了一个十分简单的初始构形例子，如图 10.6 所示，它产生了一条每 52 步就发生重复的轨道．

令初始状态是一排灰色胞腔，其余胞腔都为白色．由此出发，人们得到了一些相当美妙的图案．图 10.7～图 10.10 表明，起初这个蚂蚁的行为多么像一只蜘蛛在织一张网，但渐渐地不对称现象出现了，而大约到 9000 步时，公路的建造就开始了[①]． [65]

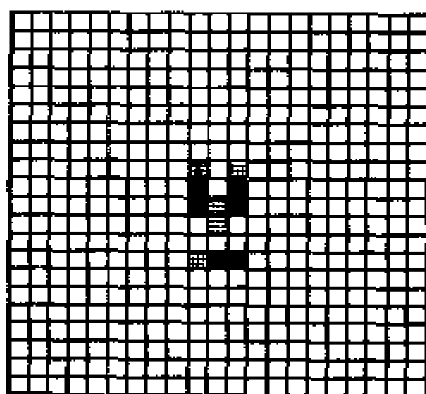


图 10.6

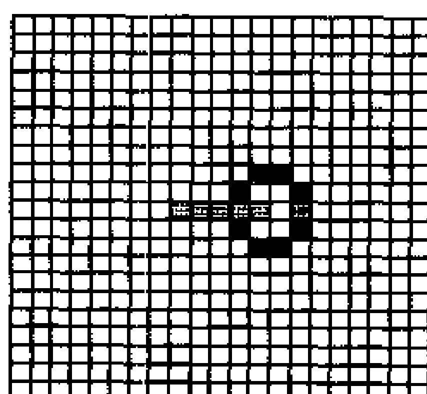


图 10.7

① 照说灰色胞腔的位置永远不变，但图 10.8 中那个孤立的灰色胞腔其位置与众不同．可能它并不是灰色胞腔，而是一个处于由黑变白(或由白变黑)过程中的胞腔；也可能是为了显示箭头(即蚂蚁)而呈灰色的．——译注

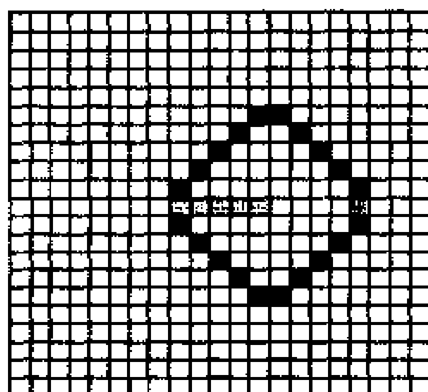


图 10.8

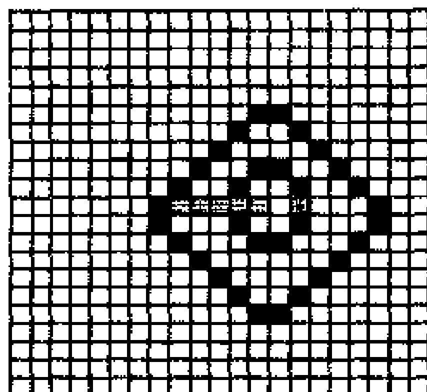


图 10.9

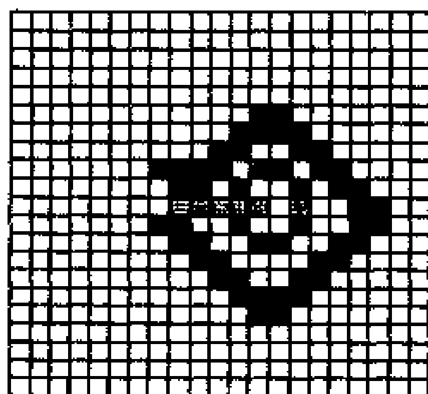


图 10.10

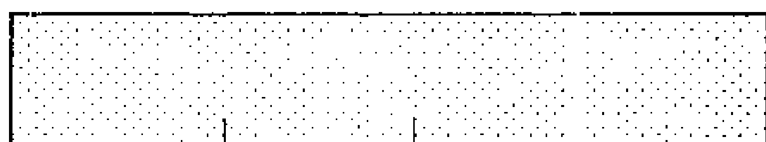
[66]

这个蚂蚁是兰顿(Chris Langton)发明的,他在某种程度上详细地研究了它的行为,包括若干个蚂蚁同时在爬的情况.

直尺作图

学习一门新学科的最好方法,正如每个人都知道的那样,是去教这门学科.有时候,如果事情顺利,人们甚至可以在最后作出原创性的贡献.最近,这种情况居然在我身上发生了,然而我却对此感到十分意外.那是我在为四年级和五年级的孩子们设计一门课程的时候.这门课程打算讲关于几何作图的内容,但并不是用传统的直尺和圆规,而是用一把"可标上记号的"直尺.

也就是说,所提供的工具只能是一支铅笔,一块橡皮擦,以及用白色卡纸裁成的一个矩形长条,人们可以在上面标上(和擦去)记号.



作为一种热身,计划要求学生们用这些工具决定在图10.11~图10.13中哪条线段比较长.作为一种变化,接下来可以要求他们在图10.14中他们认为是箭杆中点的地方标上一个点,然后用这把直尺来确定他们的估计相对于真正的中点来说是偏

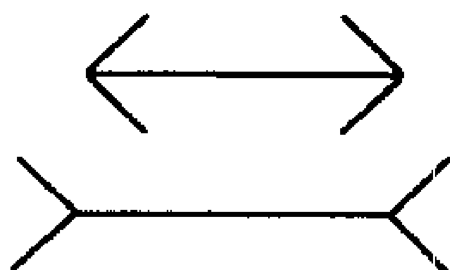


图 10.11

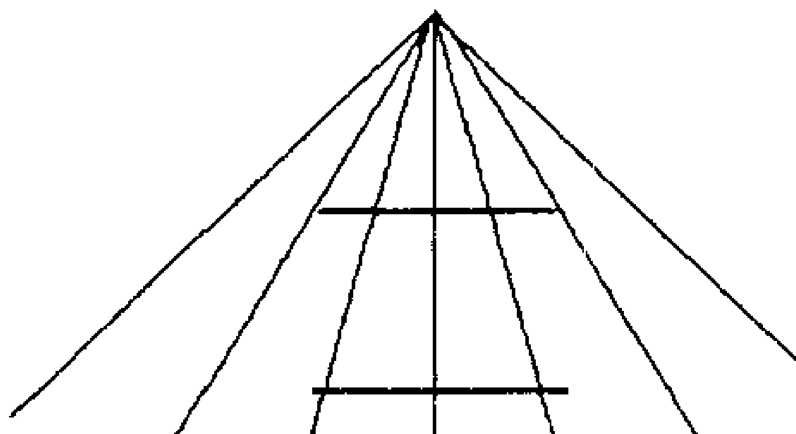


图 10.12

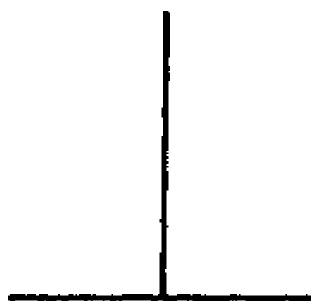


图 10.13

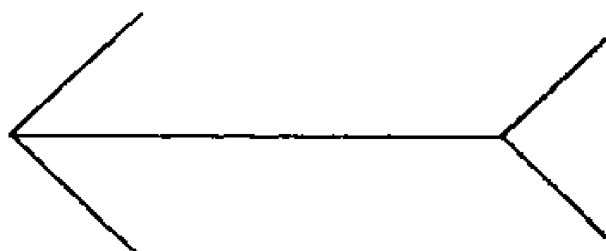


图 10.14

左了还是偏右了. 这里的想法是,看一看这些学生是否能够在这把直尺上适当地标上记号,并用来比较他们所估计的点与那箭杆两端点的距离. 这里用了众所周知的视错觉现象,目的是让他们相信,用心进行这些测量还是有一定意义的.

这后一个练习引出了作图问题的第一个例子. 给出一条线段,我们怎样才能用这把直尺来找出它真正的中点? 我所建议的解答可能将不得不演示给学生们,它是:在这条线段的两个端点处作出两条等长的垂线,然后把它们的端点连起来,如图 10.15 所示. (到这一步,我假定这把直尺是一个真正的矩形,所以作垂线仅是举手之劳. 后面我将回到这一点.) 这可以引出某些有意义的讨论. 要求学生解释他们怎么知道这样的作图确实是找出了那个中点,就把他们带进了对称这个重要的主题. (这里说几句离题的话. 可以把规则改变一下,即允许学生利用一

张普通的纸,再加上他们的铅笔,来找出这个中点,看一看有多少学生能想到在那张纸上适当地标上记号并将之折叠的方法.这也是很有趣的.据我有限的经验,一个孩子完全有可能像一个大人那样完成这件事.)

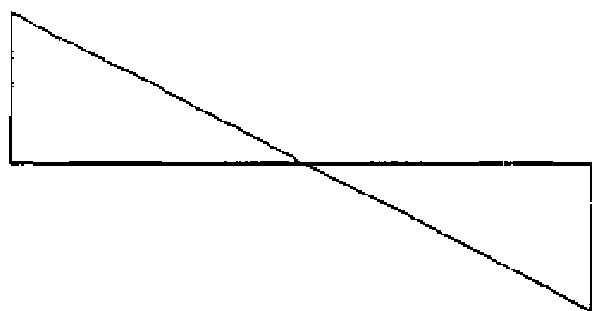


图 10.15

[68]

下一步是用图 10.16 中的作图来把一条线段三等分.这里又有一次机会引出某些相关的讨论,例如为什么我们认为右边线段的长度是整条线段长度的三分之一.在这以后,可以要求学生自己动手把一条线段五等分,等等.

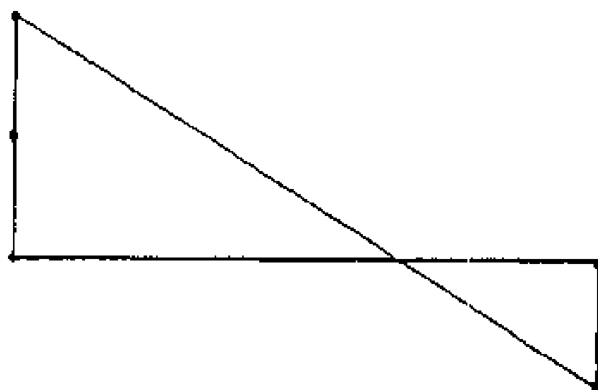


图 10.16

现在我们可以做各种各样的事了.例如有一个美妙的练习:在一张纸上给出一条水平的线段,要求学生以这条线段为底

边作一个等边三角形．这里的思路是，先作这条线段的中垂线，然后把这条线段的端点标记在直尺上(图 10.17)．

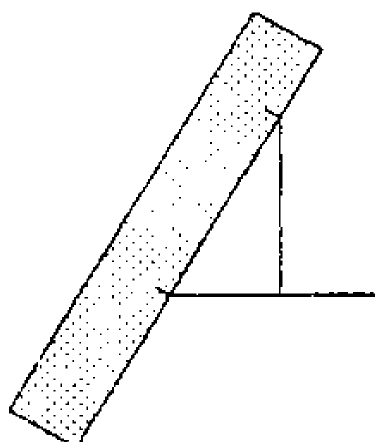


图 10.17

关于用这把标上记号的直尺允许进行哪些操作，现在有必要作出稍微准确一点的说明了．正如上面这个练习所表明的，人们可以找出一条给定直线与一个具有给定圆心和给定半径的圆的交点，尽管人们不能画出这个圆．用另一种方式叙述，给出一条直线 L 和一个不在 L 上的点 P，并在那把直尺上给出两个标记 A 和 B，人们就可以把直尺放置得使 A 落在 P 上，使 B 处在 L 上，从而在 L 上得到一个新的点 Q，当然，假定这种放置是可以实现的．实际上，所有经典的尺规作图都是可以实现的，因为从 [69] 这里不难看出作出"线段的平方根"的方法．

下一个练习是演示怎样平分一个角(在角的两条边上分别标上与角顶点等距离的点，在这些点处作出垂线，连接垂线的交点与角顶点)．我正想写下一些评注，大意是：虽然如我们前面所看到的，我们可以用我们的直尺"三等分"一条线段，但是对角来说，类似的作图已被证明是不可能的．我马上就意识到这个结论是不对的！这里有一个三等分角的作图，它归功于帕普斯(Pappus)，其中只用到一把标上记号的直尺(图 10.18)．

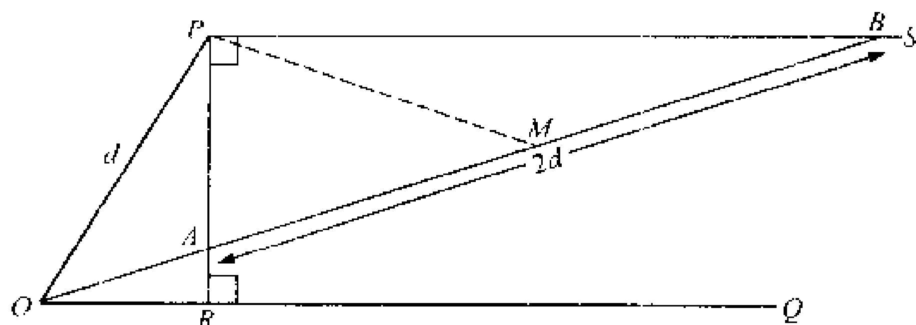


图 10.18

角 BOQ 是角 POQ 的三分之一.

证明　如图所示,线段 AB 的长度是 OP 的两倍. 现在从 P 向 AB 的中点 M 画一条直线. 那么 PM 的长度为 d,这是因为——你从这里出发便可把它拿下.

于是问题来了:什么是一把标上记号的直尺能做而一套传统的直尺和圆规不能做的事? 在上述作图中,先是任意取了点 P,然后从 P 向 OQ 作垂线. 现在在那把直尺上标上间距为 $2d$ 单位的点 A 和 B,而且我们得把直尺放置得(1) 使它经过点 O,(2) 使标记 A 处在 PR 上,(3) 使标记 B 处在 PS 上. 这是一步至关重要的动作. 更一般地,给出一条直线 L 和一个不在 L 上的点 P,以及直尺上的两个标记 X 和 Y. 现在,当把直尺放置得总是经过 P 而使 X 总是处在 L 上时,考虑由 Y 描出的轨迹. 这条轨迹称为尼科米迪斯(Nicomedes) 的蚌线,是由一个 4 次方程给出的. 图 10.19 表明了它的典型图象. 当然不允许画出这条蚌线,但我们可以找出它与一条直线的交点,就像在帕普斯的作图中那样. 在操作上,这相当于把直尺放置得使那两个给定的标记 X 和 Y 处在一对给定的直线上,然后"滑动" 直尺,始终让标记处在直线上,直到直尺经过一个给定的点. 一个很容易执行的操作.

现在有一件有趣的事,即这个蚌线动作让人们不但能三等

分角,而且事实上能作出任何最高达4次的多项式的实根或复根. 这里的技巧是证明可以作出正数的立方根. 图10.20表明了怎样做到这一点. 结论是:$\mathbf{x}$ 是 $\mathbf{a}$ 的立方根(这里用黑体字母 [70] 表示长度). 当然,这可以用解析的方法来证明,但是如果人们

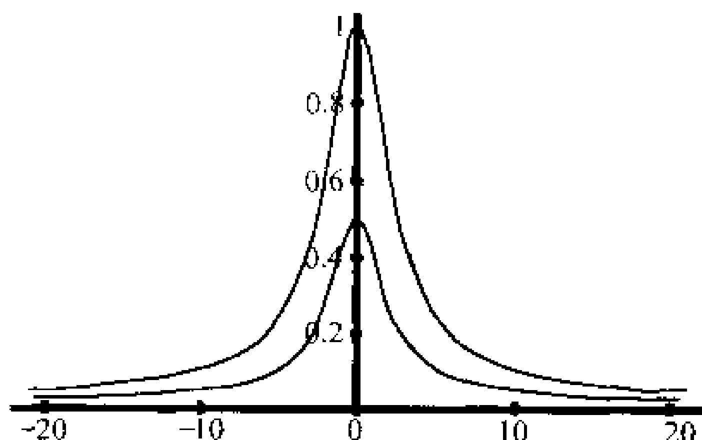


图 10.19

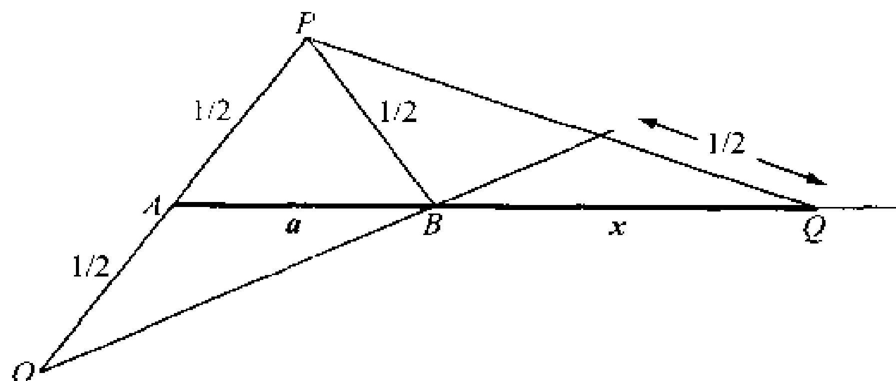


图 10.20

不把事情处理得比较合宜的话,那么计算起来就会乱成一团. 用米尼劳斯(Menelaus)定理[①]可以给出一个巧妙的证明.

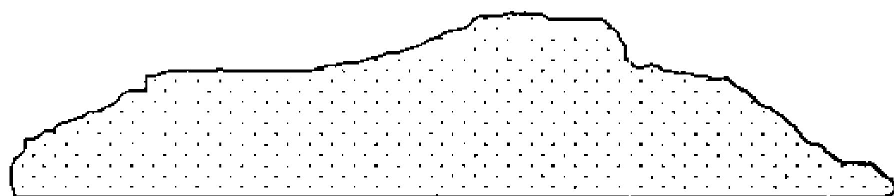
① 关于米尼劳斯定理,可参见本译丛中《近代欧氏几何学》的第8章,但那里译成"门奈劳斯定理".—— 译注

在作图时，要选定一个单位长度，在这里就是 OP. 现在画出长度为 a 的线段 AB，作出等腰三角形 APB，点 O，以及线段 AB 和 OB 的延长线. 然后利用这两条直线以及点 P，借助直尺上标好的相距 1/2 的记号，通过"蚌线动作"，定出点 Q.

最后，利用三等分角和立方根的作法，我们就可以作出复数的立方根（棣莫弗（de Moivre）定理）. 再加上平方根的作法，就足以作出任何最高达 4 次的多项式的根.

这一切的用意在于前面提到的观点，即用这种看来原始的工具，人们可以完成用经典的尺规作图操作不可能完成的作图. 例如，根据伽罗瓦（Galois）理论，我们知道可以作出一个正七边形，同样也可以作出正 13 边形和正 19 边形，但不能作出正 11 边形. 一般地说，对于素数 p，如果 $p-1$ 只有 2 和 3 作为素因数，我们就可以作出一个正 p 边形. [71]

现在回到作垂线这个题目. 人们可能要问，是不是那把直尺一定要本身连带有直角. 例如，假设它像这样：



只用这样一把直尺，我们就能作出垂线？答案是能. 要对一条给定直线在给定点上作出一条垂线 …… 但是转念一想，还是让我把这个作图问题留作练习吧. 我的解答包括画两条辅助线和两个辅助点，还要在直尺上做两个标记. 另一方面，从一个给定点向一条给定直线伸一条垂线，那得费我四条直线和三个直尺标记，但是没有附加的点. 作为一种进一步的变化，可假定我们要经过 A 和 B 两点画一条直线，但是这两点的距离大于直尺的长度. 对此我将假设，通过小心地选定一个目标，人们能够从 A 画出一条直线，它与 B 的距离不超过直尺的长度. 于是我

得到了一个多少有点笨拙的作图.

不用说,可能除了上一段中这些非常初等的事实外,前面所说的结果没有一个是新的. 这些材料的一个来源是一本由丹麦几何学家耶尔姆斯列夫(J. Hjelmslev)写的有趣的书,叫做《几何实验》(*Geometriske Eksperimenter*),它已被译成德语,作为《数学和自然科学教学杂志》(*Zeitschrift für Mathematischen und Naturwissenschaftlichen Unterricht*)的一本附册于1915年出版. 据耶尔姆斯列夫的说法,上面给出的立方根作法没有证明地出现在牛顿(Newton)的《广义算术》(*Arithmetica Universalis*, Cambridge, 1707)中. 但耶尔姆斯列夫声称牛顿的证明只是在尼科米迪斯给出的证明上稍作修改而已.

我想人们可以把这些内容再向前推进,考虑同时用一把有标记的直尺和一把圆规能够完成的作图. 这将产生最高达8次的轨迹. 或许这样的研究已经有人在做了. [72]

第11章 游戏：实的，复的，虚的

人们在玩的游戏

对策论这门学科如今已是数学的一个十分坚挺的分支，专门的著作有几十本，研究性论文有几百篇，定期出版的杂志至少有两种．对策论以许多迥然不同的面貌出现，包括经济的，政治的，军事的，当然，还有娱乐的；但所有这些表现都有一个共同之处．除了极少数例子外，用作分析的对策游戏并不是传统上已经存在的游戏，而是一些新游戏，发明这些新游戏不是供人们玩的，而是供人们分析的．有少数几个例外，如"连点成盒"（Dots and Boxes）游戏[①]，它在对策论产生之前就有了．就在最近，伯利坎普（Elwyn Berlekamp）又把组合对策论中一些最为成熟的思想用到了围棋的某些对阵局面上．然而，这个工作与实际下围棋没有什么关系，就像国际象棋中排局与下棋没有什么关系一样．

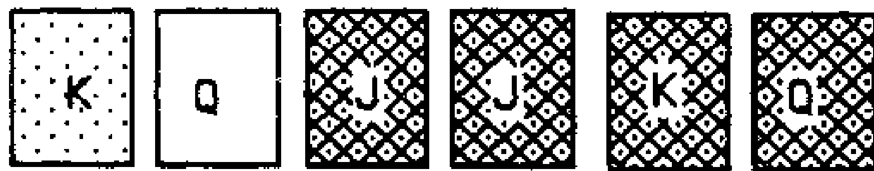
说这些话是为了引出以下这件事情：与上述说法相反，佩特

---

① 这个游戏据说源自东方．在一张画有格点的纸上，对局双方轮流把相临的点用线段连起来，如果谁连上一条线段后能围成一个（或几个）1×1的小方格（即"盒子"），谁就赢得这个（些）"盒子"，并有权再连一条线段；否则，就让对方连．待所有相临点都连上了线段，谁赢的"盒子"多，谁就获胜．可参见《萨姆·劳埃德的数学趣题》（马丁·加德纳选编，陈为蓬译，上海科技教育出版社，1999年版）中的"画盒者趣题"．近年来，人们发现这个游戏的数学内涵十分丰富，特别是可以成为组合对策论中的经典实例．——译注

森(Michael Paterson)和兹维克(Uri Zwick)最近成功地对人们熟知的"记忆翻牌"(Concentration)游戏作了几近完整的分析. 这个游戏许多读者无疑在孩童时代玩过,要不,后来也许同孩子们玩过. 一副由 n 对牌组成的纸牌,经洗牌后正面向下地摊在一张桌子上. 一位局中人翻开一张牌,接着再翻开一张牌. 如果这两张牌配成一对,这位局中人就把这对牌归己,并有权再翻两张牌. 如果不是一对,这两张选出的牌就要翻成仍然正面向下,接着便轮到对手翻牌. 最后谁手中积累的牌对多,谁就是赢家. 因为对局双方都看到了所有翻开过的牌,所以这个游戏的取胜之道在于能记住先前翻开过的牌的位置. 于是,具有照相机般记忆力的人应该玩得很好. 但是,如果对局双方都有完美的记忆力那会怎样? 这样会不会成为一件只是靠抽牌运气的事情? 这两位作者的第一个惊人结论是:事实上,可以采取有策略的行动,而且有时候故意走一步"傻着"是有利的. 我们即予说明.

考虑一副由四对牌组成的纸牌,比方说,一对 A,一对 K,一对 Q 和一对 J. 你的对手首先翻牌,他翻出了一张 A 和一张 K. 接下来你翻出了另一张 A,于是把这对 A 拿下,继续再翻. 这次你第一张牌翻出了 Q,至此游戏的局面如下图所示.



那张 K 上画了点,表示虽然它正面向下,但它的位置已为你们双方所知. 事实上,既然局中人都具有完美的记忆力,我们不妨假设翻开过的牌仍然保持正面向上. 那张 Q 本来翻开着,而其余牌的情况则不得而知. 现在你必须翻第二张牌,所以很自然,你应该去翻开一张正面向下的未知牌. 错了! 正确的做法是翻开那张 K,虽然你知道这样你将一无所获,而且要把翻牌权让给你的对手. 理由何在? 好吧,如果你去翻另外一张牌,那

么那张牌是J的概率为1/2. 在这种情况下,你的对手将把那三对牌全部拿去. 因此你的平均损失是3/2. 另一方面,如果你没有抽到J,那么你将以同等的可能性抽到K或Q,而根据对称性,可知你在前一情况下的平均损失等于你在后一情况下的平均所得,因此你的总体平均损失是3/2. 另一方面,假设你去翻开那张你知道是K的牌. 现在你的对手以1/2的概率抽得J,接下来将根据他是否抽到另一张J而赢得或失去那三对牌. 因此他的数学期望是 $-1/2$. 如果他抽到K或Q,那么他将拿下一对牌,接下来他赢得那两对牌的可能性是他失去它们的可能性的两倍. 这使他的数学期望为5/6. 因此他的总体平均所得,也就是你的总体平均损失,是1/3. 总而言之,如果你们约定一对牌以一美元计,那么你的"傻着"只使你破费33美分,而如果你按通常的方法玩,你就要破费一个半美元(从现在开始,在所有的分析中都假定你感兴趣的只是使你的平均赢得达到最大. 而你的赢得就是你所得的牌对数与对方所得的牌对数之差).

当然,上述计算背后的常识性道理是十分清楚的. 在决定是否要翻开第二张牌的时候,你在你得到一对牌的机会和相反假如你没有成功你给对方的信息之间作了权衡. 作为这方面的一个最简单的例子,考虑一副只有两对牌的纸牌. 那么先翻牌显然是不利的,因为你只有三分之一的机会得到一对牌,而如果你没能得到,这两对牌你都会失去. 那么是不是一副牌无论多少总是先翻不利? 对于一副只有3,4或5对牌的纸牌,答案是肯定的,但如果有6和7对牌,先翻牌的局中人将占得上风. 从这以后,将按这副牌所含的牌对数是奇数还是偶数而分别对先翻者和后翻者有利——一个远不明显的事实. [74]

回到有策略的玩法这个问题. 一位局中人的最佳策略依赖于他对他的对手所作的假定. 在前面的例子中,我们考虑的是一位"天真的"的对手,也就是说,一位当他的牌不能同一张翻开的牌配对时总是再去翻一张牌的对手. 我们将把这种走法称为

2－着(2-move). 在这种情况下,结论是我们差不多总是应该出那个傻着,从现在起我们把那个傻着称为 1－着(1-move). 然而,在某些情况下,我们应该走一步"超级傻着". 例如,在有四对牌的游戏中,如果你的对手开头翻出了 A 和 K,你能做的最聪明的事是把它们再翻一次. 当然这相当于放弃,然而,就像在上一段中看到的,这可能确实是最好的做法了. 我们将把它称为 0－着(0-move). 还有,由于我们假设局中人具有完美的记忆力,因此我们同样可以假设翻开过的牌将保持正面向上.

n =	
1	22
2	022
3	1022
4	21002
5	021002
6	1021102
7	21001102
8	021101102
9	2101101102
10	02101101102
11	102101111102
12	2101111111102
13	02101111111102
14	102111111111102
15	2110111111111102
16	02111111111111102
17	210111111111111102
18	0210111111111111102
19	10111111111111111102
20	210111111111111111102
21	0211111111111111111102
22	10111111111111111112102
23	211111111111111111112102
24	0211111111111111111112102
25	21111111111111111111112102
26	211111111111111111111112102
27	1011111111111111111111112102
28	21111111111111111111111112102
29	021111111111111111111111112102
30	2111111111111111111111111112102
31	21111111111111111111111111112102
32	211111111111111111111111111112102
33	2111111111111111111111111111112102
34	21111111111111111111111111111112102
35	211111111111111111111111111111112102
36	2111111111111111111111111111111112102
37	21111111111111111111111111111111112102
38	211111111111111111111111111111111110102
39	2111111111111111111111111111111111110102
40	21111111111111111111111111111111111110102

[75] 第 n 行的第 k 个数字表示当有 n 对牌而有 $k-1$ 张牌正面向上时的走法

对付一位天真的局中人的一般最优策略结果如下:除了在牌对数小于30时有个别的例外,当牌对数与正面向上的牌数具有相同的奇偶性时,我们总是应该走一步1-着. 当奇偶性不相同时,我们仍然应该走一步1-着,直到正面向上的牌占据很大的比例. 到某个临界点,在有些情况下我们突然走一步2-着,接着是0-着. 这幅图景由上页的表格给出,其中假设即使在游戏刚开始一张牌也没有翻开时,走0-着和1-着也是允许的[①].

佩特森和兹维克的工作所关注的并不是一位天真的对手,而是两位同样老练的局中人. 在这种情况下,最优策略描述起来甚至更加容易. 除了六对牌的情况是例外,规律是:如果牌对数与正面向上的牌数具有相同的奇偶性,就应该走一步1-着;否则,就根据正面向上的牌数是大于还是小于牌对数的三分之二来决定是走一步2-着还是0-着. 当然,无论哪一位局中人走出了一步0-着,就意味着这游戏即将结束,因为如果不是这样的话,对局双方就会永远不停地走0-着.

关于佩特森和兹维克这个结果的故事很吸引人. 首先,最优策略的模式原来是如此容易描述,有这种机遇真是很幸福. 当然,这个模式如果没有高速计算机的帮助是不可能被发现的(肯定没有人能事先猜到这样一种模式). 另一方面,在这种计算总会有结果的前提下,得出人们需要多少就多少的数据是一件相对平常的事. 例如,对付一个给定的有些牌正面向上有些牌正面向下的局面的最优走法,依赖于对付那些正面向下的牌较少的局面的最优走法,循此不难求出适当的递归关系. 然而,令人十分敬佩的是,这两位作者无懈可击地证明了上述策略确实是最优的. 况且,这个证明极其复杂,它分成七节,大约用了

① 若要仔细考察这张表,还请注意:1-着是指翻开一张未被翻过的牌之后,再翻一张翻过的但不能与刚翻的牌配对的牌;而再翻一张翻过的但能配对的牌则属于2-着. 另外,由于已翻开且配对的牌已经取走,因而不计在牌对数中. ——译注

十四页．或许最值得一提的是，这个证明大量地用到了计算机辅助的符号计算．用作者的话来说，“如果不借助实验和本质地使用自动化符号计算，这个分析是否能够完成是令人怀疑的．”

不应该用上述评论来说明整个研究都是由计算机操纵的，而研究者要做的事只是输入数据而已．相反，在攻克这个问题的过程中，主要的挑战在于把它变成一种可以使用决定性符号计算的形式．一些节的标题，如“算符记号”(Operator Notation)、“自举”(Bootstrapping)、“边界层影响”(Boundary Layer Influence)，指明了在这一努力中所涉及的一些概念．于是这里又有了一个例子，它使得数学家可能要重新思考一下在高超计算机时代他们关于“做数学”是指什么的观念．

与此同时，虽然有这些新结果，孩子们和家长们无疑会继续把“记忆翻牌”玩下去，因为不管是好是歹，看来人类没有迹象会发展到具有完美的记忆力．

人们不玩的游戏

这里有一个不同类型的记忆游戏．我有一副有着无穷多张牌的纸牌，我第一着把它们的一个有限子集给你．你可以扔掉[76]一张牌，然后我把我手中余下牌的另一个有限子集给你．你仍可扔掉一张牌．这个游戏就这样连续地玩了可数无穷多步．最后如果你手中空无一牌，你就赢了，否则我赢．就这么简单[①]．

你看，上面所描述的这个游戏不是很有趣，因为事实上你有一个很容易的必胜策略．你只要在我给你牌的时候把这些牌记清楚，并不断地扔牌，首先扔我第一着给你的那些牌(以任何次

① 这个游戏乍看似不好理解．事实上，“我”一共给了“你”可数个不交的有限集，它们的并是一个可数集；而“你”一共可以扔掉可数张不同的牌，这些扔掉的牌也组成一个可数集．这两个可数集之间一定存在一一映射．问题就是要找出一个恰当的一一映射，使得“你”每次扔掉的牌都属于“我”已经给你的牌．这里的有关概念，可参见本译丛中《无穷之旅》的第9章．——译注

序),其次扔我第二着给你的那些牌,如此下去.于是,到这个游戏结束时,你就把它们扔得一干二净了.但请注意,要执行这个策略,你必须具有非常强的长期记忆力.例如,如果我每一着给你一百万张牌,你会越来越落后,到你把我第一手给你的牌全扔完的时候,你手中将握有 10^{12} 张牌[①].换句话说,要玩这个简单的策略,你必须具有一种无限的记忆能力.因此,让我们走另一个极端,假定你一点记性也没有,只是在每个轮次你知道你手中所握牌的名称或点数(请注意这类信息对于如国际象棋、西洋跳棋[②]、"连城"[③]之类的有限对策是十分足够了,在这类游戏中,要决定下一着如何走,有关系的事仅是当前的盘面情况,而不是这个盘面是怎么造成的).如果这副纸牌是可数的,你的情况仍然很好.在游戏开始前,你以某种方式给出自然数集与这副纸牌的一个一一映射,然后在每个轮次你只要扔掉手中映成最小数的那张牌就行了.然而,如果这副纸牌是不可数的又会怎样呢?

定理 1 *对于零记忆游戏中的一副不可数的纸牌,第二位局中人没有必胜策略.*

证明 无记忆游戏中的一个策略只不过是一个函数,它把纸牌的每一个有限子集(你手中的牌)映成一张牌(你决定扔的牌).因此,特别地,它把每一对牌映成你决定要扔的一张牌.于是让我们对每张牌 x 定义一个集合 D_x,它是由所有这样的牌 y 组成的:$y \neq x$,而且当你手中握有 $\{x, y\}$ 时你将扔掉 y.现

---

① 严格地说,应该是 $10^{12}-10^6$ 张牌.当然,这无关紧要.——译注

② 西方流行的一种棋类游戏.棋盘同国际象棋棋盘.对局双方各执12子,均放在己方底部三行的黑格内.棋子可沿对角线走至左(右)前方空格内.若该空格已被一对方子占据,而再左(右)前方是一空格,则可越过对方子跳入那空格,并把被越过的对方子"吃"掉.一棋子走到对方底行即升为"王".王不但可向左(右)前方行动,还可向左(右)后方行动.将对方棋子全部封死或吃掉者为赢家.——译注

③ 即对局双方轮流在一个"井"字形方格内画○和×,以先列成纵、横或对角线上的一排者为胜的游戏.——译注

在,要证明你没有必胜策略,只要证明存在某个 x',使得 $D_{x'}$ 是无穷集就足够了,因为这样它就会包含某个无穷序列 $y_1, y_2, \cdots$. 于是我第一着将给你 $\{x', y_1\}$,而在第 n 着给你 $\{y_n\}$,你就永远也扔不掉 x' 了. 为了完成证明,我们构造这样的一个 x'. 具体地说,选取纸牌的一个可数集 C,对于其中的任何元素 x,D_x 都是有限集(如果做不到这一点,我们就已经得证了). 于是 C 和所有 D_x 的并 U 显然是一个可数集. 但是这样 U 的补集(不会是空集,因为那副纸牌是不可数的)中的任何一个元素 x' 都具有性质:C 被包含在 $D_{x'}$ 中[①]. 这就是我们所期望的构造. ■

因此在零记忆游戏中你不会有必胜策略. 然而人们发现,你不需要有无限的记忆能力才能赢. 你要记住的只是你刚扔掉的那张牌(它也许正面向上地正躺在那堆扔掉的牌的顶上),而当(且仅当!)允许使用选择公理时,你还是拥有必胜策略.

定理 2 *若使用选择公理,则对第二位局中人存在一个必胜策略,条件是他既知道自己手中的牌也知道他刚扔掉的牌.*

证明 这个必胜策略十分简单. 把这副纸牌良序化[②]. 于[77]是你第一着把手中最大的牌扔掉. 此后,(a) 如果存在小于你刚扔掉的牌的牌,就把其中最大的牌扔掉. 否则,(b) 把手中最大的牌扔掉. 要明白这种做法是有效的,请注意只要你处于情况(a),刚扔掉的牌就会作为时间的一个函数呈严格递减,因此根据良序的性质,这个序列一定是有限的. 所以情况(b)迟早会发

---

① 要证明这一点,请注意对 $x \in C$,$\{x, x'\}$ 中总有一张牌要被扔掉. —— 译注

② 粗略地说,良序是一个集合中元素间的这样一种大小关系:在这个大小关系下,这个集合的任何一个非空子集都有一个最小元素. 具有良序的集合称为良序集. 例如,自然数集按通常的大小关系就是一个良序集(请回忆第8章提到的佩亚诺归纳公理). 用选择公理可以证明,任何一个集合都可成为一个良序集,此即所谓良序化. 这个结论称为良序定理. 本章除了用到良序定理外,还将用到结论"良序集中的严格递减序列必为有限序列". 详情可参见《超穷数与超穷论法》(谢邦杰编著,吉林人民出版社,1979年版). —— 译注

生. 现在考察你手中的某张牌 x. 如果它小于当前扔掉的牌,那么根据上面的论述,你一定会在未来的某个时刻把它扔掉. 如果它大于当前扔掉的牌,那么当情况(b) 发生时,你不把 x 扔掉也会把某张大于 x 的牌扔掉,于是前面的论证再次适用. ■

还有第三个结果,也用到了良序,它是说:如果允许第二位局中人在走每一着时既可以看他手中的牌也可以看所有已经扔掉的牌的集合,那么他也会拥有一个必胜策略[①].

定理 1 是富勒(Martin Furer) 和斯佩克(Ernst Speck) 的一个一般结果的一个特例,而定理 2 和提及的第三个结果是由拉弗(Richard Laver) 和采谢利斯基(Krzystztof Ciesielski) 证明的. 据我所知,这些工作都没有发表过.

人们能玩的游戏

下面这个被马丁·加德纳戏称为"大嘴巴"(Chomp) 的游戏,是 Nim[②③]的一个变种. 它曾流行了好一阵子,但是最近得到的一些结果引出了许多有趣的问题. 这个游戏很容易描述. 先用"甜饼干" 摆出一个矩形阵列. 接着局中人开始轮流走子,走法是选取一块饼干,再把这块饼干右上方的所有饼干都拿走. 图 11.1 显示了一个 3 × 5"大嘴巴" 游戏开始时的局面和第一位局中人首先在第三行第四列处[④]作了选取后的局面.

被迫捡起左下角那块毒饼干的局中人就是输家. 这个游戏的本质特征是,只要采取最优玩法,应该总是第一位局中人赢.

---

① 这个结果与定理 2 的区别看来是:虽然可以看到所有扔掉的牌,但不知道哪一张牌是刚扔掉的. 请注意这第二位局中人一点记性也没有. —— 译注

② 参见附录 1. —— 原注

③ 此游戏据说源于我国,玩法是:对局双方轮流从若干堆棋子中取子,每次只能在一堆中取,数目不限,但至少一个,以最后取子者为输(或赢). 可用二进制数算得这个游戏的必胜策略. 可参见本译丛中《数学游戏与欣赏》的第 1 章. —— 译注

④ 在这里,行序是从下到上,列序则仍为从左到右. —— 译注

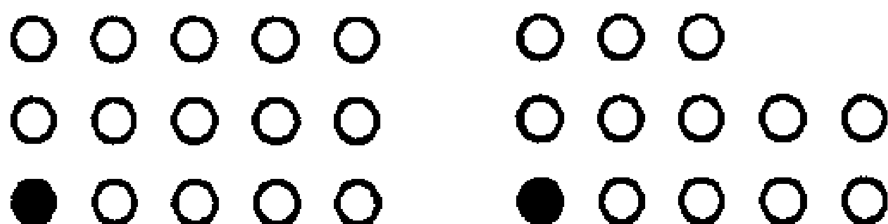


图 11.1

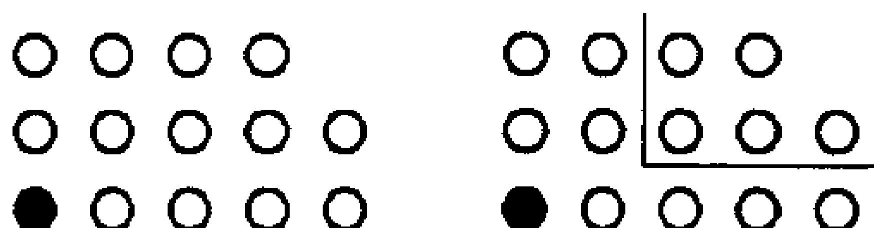


图 11.2

[78] 这个证明可用图 11.2 表示.

如果拿走右上角那块饼干(如图 11.2 左)可以给第一位局中人一个必胜局面,那么结论已得证. 如果不是这样,它就给出了一个必败局面①,于是第二位局中人就有一个应着使自己得到一个必胜局面(如图 11.2 右);但第二位局中人的任何这样一着都可被第一位局中人在自己走第一着时采用.

应该指出,这个论证完全是非构造性的,它只断言对第一位局中人存在一个必胜策略,但对怎样找到这个策略毫无启迪. 确实,除了两个特例外,这些必胜策略的构造如何,人们几乎一无所知. 一个特例就是 $2\times n$(和 $n\times 2$)的游戏. 容易看出,第一位局中人总能给对方一个下行比上行多一个的局面(图 11.3):

---

① 这里用到:对于这种游戏,要么第一位局中人有必胜策略,要么第二位局中人有必胜策略. 其根据是著名的策梅洛(Zermelo)定理. 可参见《对策论(博弈论)讲义》(中国科学院数学研究所第二室编,人民教育出版社,1960 年版). 另,"局面"原文为 position. 一般译"阵地"或"位置",但译者以为还是"局面"为好.——译注

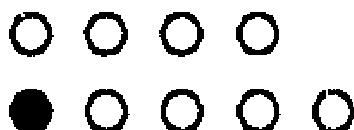


图 11.3

因此他最终会赢得胜利. 另一个例子是 $n \times n$ 的情况. 这里第一位局中人可先采取如图 11.4 所示的走法,然后采取与他对手的走法以对角线为对称的走法.

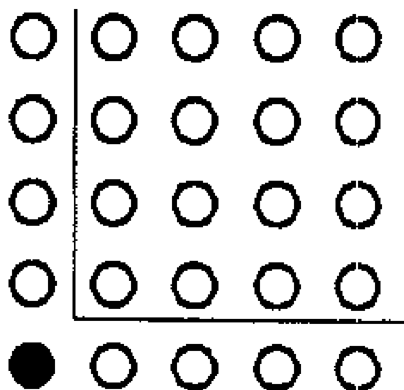


图 11.4

我们的愚昧无知可用下例说明. 从实验看,$3 \times n$"大嘴巴"的开局胜着似乎是唯一的. 对 3×5 的胜着如图 11.1 右所示. 但这些开局胜着没有明显的模式. 起初猜想开局胜着总是唯一的,结果却发现 6×13 的游戏是一个反例,它有两个开局胜着.

"大嘴巴"也能用无穷多块甜饼干来玩. 最简单的例子是 $2 \times \omega$[1]的情况(图 11.5). 请注意,即使有着无穷多块饼干,这个

[1] ω 是自然数集的序数. 这里对序数仅作一简介. 序数是本质上相同的良序集的一个共同特征. 有限序数就是 0 和自然数,但还有超穷序数,如 ω. 对序数可进行加、乘、取幂等算术运算,但运算律有其特别之处,如加法不遵从交换律. 若干序数组成的集合在一个自然的大小关系下成一良序集. 可数良序集的序数称为可数序数. 若干序数总有一个上确界,可数个可数序数的上确界仍为可数序数. 详情可参见《超穷数与超穷论法》(谢邦杰编著,吉林人民出版社,1979 年版). —— 译注

游戏仍然可以只用有限多步完成．事实上，这是一种应该总是[79]第二位局中人赢的游戏的第一个例子．容易看出，不管第一位局中人怎么走，第二位局中人总能走出如图 11.3 的局面．于是可以推出，如图 11.6 所示的 $\omega \times \omega$ 的游戏应该总是第一位局中人赢．通过选取第三行第一列的那个元素，他就能给对方一个 $2 \times \omega$ 的游戏在开始时的局面．

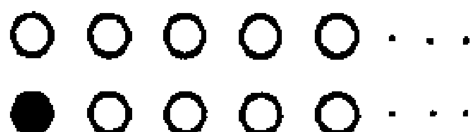


图 11.5

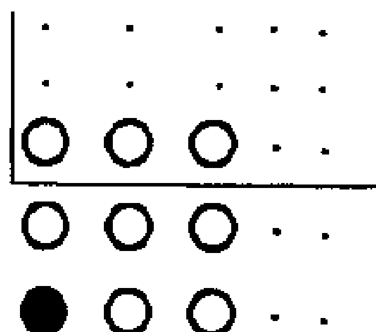


图 11.6

用一种显然的方式，这个 $\omega \times \omega$ 的游戏可以表示为平面上第一象限的格点集合，而且这可以推广到三维或更高维的情况．在三维空间第一卦限上的这个游戏仍然悬而未决；这是一个应该总是第一位局中人赢的游戏还是一个应该总是第二位局中人赢的游戏？先前提到的那些新结果把这个问题归约成超穷序数算术中的一道题目，下面即予说明．

首先，我们指出，"大嘴巴"可以在任意序数上玩；例如，给出任何两个序数 α 和 β，我们就可以考虑 $\alpha \times \beta$"大嘴巴"．因为

根据序数的良序性质,可以推出这些游戏的任何对局一定会在有限步内结束.类似地,人们还可以考虑在三个或更多个序数的笛卡儿积上玩的游戏.

其次,对一位正要走子的局中人来说是必败的一个游戏局面已被伯利坎普、康韦和盖伊称为P局面(P-position)(见《取胜之道》(*Winning Ways*),Academic Press,1982),这里P表示这个游戏应该总是那位在这之前刚走了一着的局中人(previous mover)赢.图11.3,图11.4和图11.5都是"大嘴巴"的P局面.下面这件事是赫德尔斯顿(Scott Huddleston)告诉我的.它是一个定理的一种特殊情况,那个定理他建议应该叫做"大嘴巴基本定理".我们陈述如下:

定理 对于任何正整数n,存在一个唯一的序数α,使得$n \times \alpha$矩形阵列是一个P局面.

证明 我们首先指出,这个临界序数的唯一性唾手可证.如果α给出了一个P局面,而$\alpha < \beta$,则$n \times \beta$就不是一个P局面,因为第一位局中人能够把局面走得变成$n \times \alpha$. [80]

我们已经看到,如果$n = 2$,则$\alpha = \omega$.赫德尔斯顿确信矩形阵列$3 \times \omega^{\omega}$和$4 \times \omega^2$是P局面,但是我还没有看到这些有趣结论的证明,他也没有对那个"基本定理"提供一个证明.但是伯格曼倒提出了一个证明,下面就是对他的论证稍作修改而得到的一个证明.事实上,我们将证明,这个临界序数α一定是一个可数序数.为此,从现在开始,这里考虑的所有序数都假设是可数的①.

我们不用甜饼干,而是把"大嘴巴"的一个局面看作一个非增的序数数列$(\alpha_1,\cdots,\alpha_n)$,这样更好一些.对任意的正整数

---

① 在结论没有被证明之前,就根据这个结论作出某种假设.这种说法有点令人费解.事实上,在下面的证明中,我们可以在定义$f(\alpha,m) = \beta$时把α和β限定为可数序数.——译注

$m(0 < m < n)$ 和序数 α①,如果存在一个序数 β,使得 $(\beta,\beta,\cdots,\beta,\alpha,\alpha,\cdots,\alpha)$(其中 β 出现 m 次)是一个P局面(这是在一个 $n\times\beta$ 的游戏中第一着把上面 $n-m$ 行的 β 用 α 代替后而形成的局面②),我们就定义 $f(\alpha,m)$ 为这个唯一的 β③. 令 $g(\alpha)=\sup f(\xi,m)$,取上确界的范围是所有的 $\xi<\alpha$④和 $0<m<n$. 因为这些 (ξ,m) 组成的集合是可数的,所以 $g(\alpha)$ 是一个可数序数. 现在定义序数序列 (α_i),其中 $\alpha_1=1$,α_{i+1}⑤ $=g(\alpha_i)$,并令 γ 是大于所有 α_i 的最小序数. 如果 $n\times\gamma$ 是一个P局面,那么定理已得证. 如果不是,那么存在一个开局胜着,它给出一个P局面 $(\gamma,\cdots,\gamma,\xi,\xi,\cdots,\xi)$,$\xi<\gamma$,其中 γ 出现 m 次. 让我们看看是否可能有 $m>0$. 如果是那样的话,那么 γ 就会是 $f(\xi,m)$,而这将与 γ 的定义矛盾. 因此,那个胜着一定是给出了局面 $(\xi,\xi,\cdots,\xi)$,于是这个 $n\times\xi$ 的游戏就是一个P局面. ■ ⑥

请注意,就像对普通"大嘴巴"游戏应该总是第一位局中人赢的证明那样,这个证明完全是非构造性的,它对找到那个临界序数没有提供任何线索.

用一个稍稍复杂一点的论述可以证明,对任何一对序数 α 和 β,存在一个唯一的序数 γ,使得 $\alpha\times\beta\times\gamma$ 的游戏应该总是第

---

① 符号似有混淆,这个 α 是指任意的(可数)序数,并不是我们要证明存在的临界序数 α.——译注

② 注意这里隐含了 $\beta>\alpha$. 这一点在下面的证明中是必需的.——译注

③ 如果这样的 β 存在,则它显然是唯一的. 如果对某个 α 和某个 $m(0<m<n)$,这样的 β 不存在,一般认为 $f(\alpha,m)$ 无定义,但也可定义 $f(\alpha,m)=0$. 事实上,在公理集合论中,0就是空集 $\emptyset$.——译注

④ 似应为 $\xi\leqslant\alpha$,否则在下面的证明中推不出矛盾.——译注

⑤ 原文作 α_i+1,显然有误.——译注

⑥ 如前所述,这只是所谓"大嘴巴基本定理"的一个特例. 那么,"基本定理"本身是什么呢?看来应是:对于任何序数 α,总存在唯一的序数 β,使得 $\alpha\times\beta$ 矩形阵列是一个P局面. 事实上,下文将说到在三个序数的笛卡儿积上的"基本定理". 因此这里关于两个序数的"基本定理"又成了一个特例.——译注

二位局中人赢(因此是一个P局面).赫德尔斯顿断言 $2\times 2\times\omega^3$ 是一个P局面.如果人们知道哪个序数 α 可使得 $\omega\times\omega\times\alpha$ 的游戏成为一个P局面,这道题目就能得到解决.一个可能比较简单的问题是判定 $3\times 3\times\omega$ 的游戏是否应该总是第一位局中人赢.谁能接手?

一个现代背景下的老故事

一家传统的女子学院正在考虑接纳男生.校董会主席认为这是一个好主意,但是院长不同意."如果你接纳男生,"她说,"不到今年年底,半数以上的女生就会退学."作为一种妥协方案,他们同意只吸收1%的男生来做一次试验.

年终来临了,在董事会议上这位主席得意地宣布试验获得成功.不错,他承认,女生人数是有所减少,但是她们所占的百分比只下跌了1%,从年头的99%降到年底的98%.

"这正如我所言!"院长说. [81]

[82]

第12章 称硬币 化方为方

看低了数学

唐纳德·J·纽曼(Donald J. Newman)

在我们数学生涯的幼年时期,我们都是从一类问题上"长出乳牙"的,那就是所谓的称重问题[①]. 从中我们学到了令我们获益匪浅的"分支"过程:如果发生了如此这般的情况,我们就做如此那般的事,如果不发生,我们就反过来做如此这般的事.

这些"分支"过程在称重问题中以一种根本性的方法出现,或许这就是解决一般问题的正确方法. 我们甚至被引诱得浮想联翩. 这可能是数学中所普遍遵循的正确途径. 而如果它对数学是这样,那么它就可能是思考一切问题的正确方法. 哇!

在《纽约人》(*New Yorker*)上发表的一篇文章中,伯恩斯坦(Jeremy Bernstein)不无赞赏地把会运用这个分支推理过程认定为真正具有数学天才的一个标志. 他所举的例子是著名的12硬币问题,他所指的解答者是当时为哈佛大学本科生的策马赫(Charles (Ariel) Zemach).

[83] 啊,这个分支推理看来是极其重要而且几乎无所不能.

---

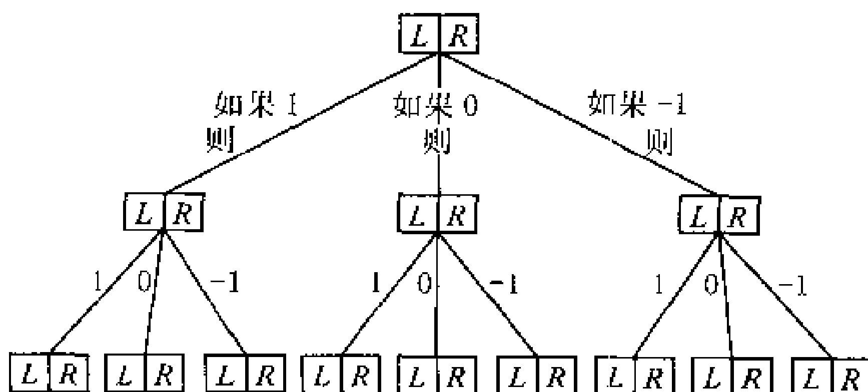
① 这是一类常见的游戏数学问题:设有一批物件,其中除了一个以外重量相同,要求逐次从中选取两组物件并对它们的总重量进行比较,以最少的比较次数找出那个重量不同的物件. 如下面将要说到的12硬币问题.——译注

然而，我们这篇文章的目的是消除这个观念！我们就用这个12硬币问题以及其他几个例子来做说明，并且希望能让读者相信，这个分支推理或许是根本不需要的．任何时候，哪儿有一个用到分支推理的解决方案，哪儿就会有另一个不用它的解决方案（而且作为一种结果，它显得更为清爽，更为简单）．

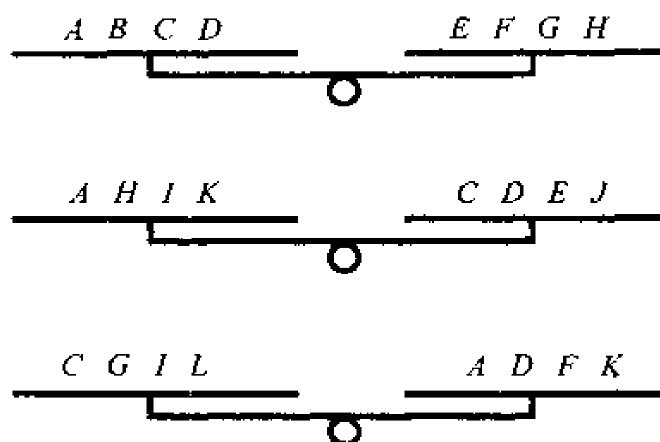
12硬币问题

给出十二枚看起来完全一样的硬币，我们已知其中有一枚硬币重量与其他11枚不同．题目是要求确定它是哪一枚硬币，它是重了一点还是轻了一点，我们只能用一台天平对这些硬币称三次．

首先请注意，即使在找出了一个解决方案之后，要对它作一番描述，看来也需要用到分支过程．一次称量的**结局**是指天平左托盘的下降、上升或保持水平．我们把这些结局分别用1，－1和0编码．每次称量包括挑选一对具有同样元素个数的子集 L 和 R，并把它们分别放在左托盘和右托盘上．一个称量过程的“流程图”就像下面这个样子：



其中对每个 L 和 R，人们都得列出相应子集的元素．请把这个过程与下面的非分支操作进行比较．把这些硬币用 $A, B, \cdots, L$ 标记，然后如下操作：



容易验证这个操作可以给出期望的结果. 具体地说,如果 A 是那枚伪币,而且它比其他硬币重,那么结局将是 1,1,-1;而如果它比其他硬币轻,结局将是 -1,-1,1. 如果 B 是那枚坏硬币,而且它比其他硬币重,那么结局将是 1,0,0;而如果它比其他硬币轻——但请等一下. 我们没有必要把所有的情况都过[84]一遍. 我们要做的只是提请注意任何两枚硬币都没有相同或相对的"历程";也就是说,任何两枚硬币都没有老是处在同一托盘上或老是处在相对的托盘上. 因此,对24种可能**状态**(即哪一枚硬币是假币,它是较重还是较轻)中的每一种,都有一个不同的结局. 这样,给出了结局,我们就知道了状态. 例如,如果我们知道了结局是 -1,0,1,那么一定是硬币 G 比其他硬币重. 顺便指出,这些称量无论按哪一种顺序进行,都不会造成差别. 还要指出的是,我们可以解决一道稍微难一点的题目,其中我们认可没有一枚硬币是假币的可能性,这种情况当且仅当结局为 0,0,0 时发生.

但现在的问题是,要设计出这三次称量,需要怎样一种灵性. 回答是——**什么也不需要**. 我们可让这个解决方案自动给出. 这里我们事实上希望传达的信息是:过去那个复杂的耍小聪明的解决方案是一种浪费精力,是一种不可取的思路!我们这个新解决方案是从27种结局出发,反推到那24种可能的假币

情况．下面是具体过程．

首先，我们作出一列 12 个不同的结局向量，要求没有一个向量与其负向量同时在此列中．要做到这一点，一个简单的方法是按字典序排出按字典序为正的向量[①]．如下表中的各列：

A	B	C	D	E	F	G	H	I	J	K	L
0	0	0	0	1	1	1	1	1	1	1	1
0	1	1	1	−1	−1	−1	0	0	0	1	1
1	−1	0	1	−1	0	1	−1	0	1	−1	0

现在，为了让那个操作过程可以进行，我们必须使得每一行中 1 和 −1 的个数相等．底行照原来样子就符合要求．倒换 C 列中的符号就把中间那行搞定，而倒换 F，H，J 和 L 列中的符号就把顶行给对付了．于是我们有

A	B	C	D	E	F	G	H	I	J	K	L
0	0	0	0	1	−1	1	−1	1	−1	1	−1[②]
0	1	−1	1	−1	1	−1	0	0	0	1	−1
1	−1	0	1	−1	0	1	1	0	−1	−1	0

现在每一行就对应着一次称量．具体地说，把那些对应 +1 的硬币放在左边而把那些对应 −1 的硬币放在右边 —— 就这么回事．进行称量，写下结局，即可从表中迅速读出那枚罪恶的硬币[③]．（表中的大写字母在排列上与先前给出的有所不同，但这显然没有关系．）

---

① 令 −1，0，1 分别为 a，b，c，于是向量（−1，−1，−1）可写作 aaa，（−1，1，0）可写作 acb，等等．字典序就是向量写作这些字母串时通常在字典中的排列顺序，而按字典序为正的向量就是排在（0，0，0）后面的向量．—— 译注

② 原文为 1，显然有误．—— 译注

③ 此表只给出了那枚假币稍重时的结局向量，而假币稍轻时的结局向量则由这些向量的负向量给出．—— 译注

在有些情况下，非分支的解决方案很容易找到．我在1至8中选定了一个整数，你必须通过问三个回答为“是”或“不”的问题把这个整数猜出来．当然，人人都会用那个分支式的或者“对话式”的不断对分策略，但是并没有必要这样做．为什么不在一开头就以任何次序只问这样三个问题：这个数在集合$\{1,2,3,4\}$中吗？在集合$\{1,2,5,6\}$中吗？在集合$\{1,3,5,7\}$中吗？显然，如果允许问的问题必须是“这个数大于x吗”的形式，这种方法就行不通了．这个例子提示了一般情况下的做法．存在一个可能状态的集合，我们想要得知真正的状态．每个问题，或每次称[85]量，或每次“实验”，都给出了这个集合的一个划分①．我们显然可以把k个划分中的集合的所有交集所组成的划分定义为这k个划分的交．那么当且仅当我们能够找出n个划分②，它们的交就是由单元素集组成的划分时，真正的状态就可以用n次实验而不用分支过程得知．

在我们下面的例子中，我们选择了两道著名的题目．其中有四枚硬币，每一枚硬币可能是真币，也可能是假币．现在我们不知道有多少是真币有多少是假币，我们必须准确地说出哪枚是真币哪枚是假币．

在一台准确的秤上称三次

对于任何一个选定的子集，我们的“秤”将准确地告诉我们其中有多少是真币（不是哪一枚是真币，而是有多少是真币）．

我们的解决方案仍是非分支的．取硬币1和2作为我们第一个子集；第二个子集，硬币2和3；第三个子集，硬币1，3和4．

这些就是我们三次“称量”的对象．我们把三次称量的结果

① 所谓一个集合A的划分，是指这个集合的一个子集族$\{A_i\}$，它满足$A=\bigcup A_i$，且当$i\neq j$时有$A_i\cap A_j=\varnothing$．——译注

② 必须指出，这里是指找出n个具体的实验分别给出这n个划分．——译注

称作 a, b 和 c. 根据它们作出确定的方法很迷人. 首先我们把这三个数加起来,得到 $a+b+c$,其中对硬币 1 计了两次,对硬币 2 两次,对硬币 3 两次,但对硬币 4 只是一次. 这样,把 $a+b+c$ 以模 2 取剩余,我们就得到了硬币 4 的本质特性. 然后把这个余数从 c 中减去,我们得到了关于硬币 1 和 2,硬币 2 和 3,硬币 1 和 3 的平衡方程组. 这就确定了前三枚硬币的性质.

对一名说谎者的七个问题

回到我们那四枚硬币,但这次我们可以就它们问任何回答为"是"或"不"的问题. 不过,我们问的那个人被允许在回答其中(至多)一个问题时说谎. 因此我们必须问得比我们问一位诚实者时所需要的那显然的四个问题要多. 事实上,使用分支过程,对这名说谎者必须提的问题的正确个数是七.

我们的目的还是用七个非分支的问题作出同样的确定. 我们开头的四个问题十分简单:1 是真币吗? 2 是真币吗? 3 是真币吗? 4 是真币吗? 下面的三个问题可能看起来用了分支过程,但是在实际上,提问的时候没有用,不过在回答的时候用了.

问题 5:你对问题 1,2 和 3 的回答是正确的吗?
问题 6:你对问题 2,3 和 4 的回答是正确的吗?
问题 7:你对问题 1,2 和 4 的回答是正确的吗?

于是确定的结果可如下得出:如果对问题 5,6 和 7 的回答中至多有一个为"不",那么开头的所有四个回答都是实话,硬币的情况就此确定. 如果对问题 5,6 和 7 的回答为"是","不","不",那么对问题 4 的回答是谎话. 所以把这个回答反过来就得到了正确的结果. 类似地,如果这些回答为"不","是,"不",只要把对问题 1 的回答反过来. 如果回答为"不","不","是",那就把回答 2 反过来. 而最后,如果它们是"不","不","不",那就把回答 3 反过来.

再说化方为方和化矩为方

戴维·盖尔(David Gale)

化方为方或化方为矩的意思是把一个正方形或一个矩形铺砌(划分)成一些子正方形的一个并集. 对于边长可公度的① 矩形, 存在着平凡的铺砌(令长为 h 而宽为 w, 如果 $h = (p/q)w$, 则用边长为 $h/p = w/q$ 的正方形的一个 $p \times q$ 阵列来铺砌). 然而,令人感兴趣的是**完美铺砌**(perfect tiling),其中任何两个子正方形尺寸都不相同. 第二个自然的限制是要求子正方形的任何一个子集都不能形成一个子矩形. 这样的铺砌称为**单的**(simple). 这里是一个很不完整的简略编年史,它只涉及了关于这个问题的部分工作.

1903. 德恩(Dehn) 证明,如果一个矩形能被一些正方形所铺砌,那么所有这些正方形的尺寸一定是可公度的.(我们说一个正方形的尺寸,意思是指它的边长.)

1925. 莫伦发现了矩形的一种用了九个正方形的完美铺砌.(我们把这些正方形的个数称为这个铺砌的阶(order).)

1939. 斯普拉格(R.Sprague) 发表了被化方的正方形的第一个例子. 它的阶是 55.

1940. 布鲁克斯、史密斯、斯通和塔特证明,矩形的完美铺砌不可能有小于 9 的阶,而 9 阶的铺砌只有两种.

1948. 威尔科克斯(Wilcocks) 发现了一个 24 阶的被完美(但不是单的) 化方的正方形.

1960. 布弗坎普(C.J.Bouwkamp) 及其合作者发现了矩形的 4094 个阶小于 16 的单铺砌(其中 3663 个是完美的,431 个是非完美的),包括 2609 个 15 阶完美铺砌.

① 即相邻两条边长之比为有理数. 一般地,一些线段的长度之比为有理数,则称这些线段是可公度的. —— 译注

1962. 杜伊韦斯京(A.W.J.Duijvestijn)证明不存在阶小于21的被完美单化方的正方形.

1978. 杜伊韦斯京发现了一个21阶的被完美单化方的正方形,并证明这是唯一具有这个阶的被完美单化方的正方形.

1992. 布弗坎普和杜伊韦斯京发表了一张带插图的"一览表",其中是从21阶到25阶的所有被完美单化方的正方形(不计明显的对称情形),一共有207种铺砌——21阶的1种,22阶的8种,23阶的12种,24阶的26种,以及25阶的160种.

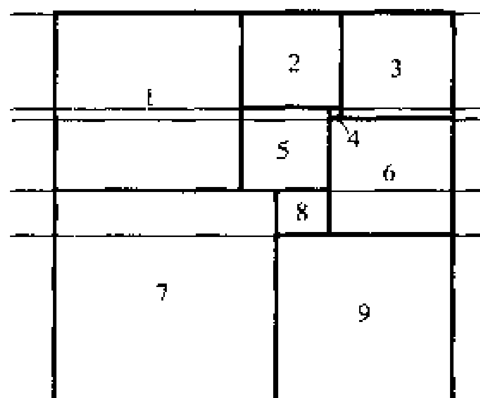
如果我们把被铺砌的正方形取作单位正方形,那么根据德恩的结果,所有子正方形的尺寸都是有理数.因此我们可以通过一种一致性的延伸予以正规化,使所有的边长成为互素的整数.这时那个大正方形的尺寸就称为这种铺砌的"既约尺寸"(reduced size).在160种25阶的铺砌中,最小的既约尺寸为147,最大的为661.

即使在30年以前,人们就知道矩形的2600多种用了15个正方形的完美铺砌,因此人们料想矩形的25阶铺砌的种数一定是非常大,特别是如果人们不要求这些铺砌是单的或完美的.布弗坎普说,(当阶大于20时)对每一个这样的正方形,大约存在着5 000 000个被完美单化方的矩形[①]!甚至可以问,人们能不能肯定,随着阶的越来越大,铺砌的种数会保持有限?回答是肯定的,但是这证明却不是十分显然.然而,它可能十分适合于在大学本科的线性代数课程中用来作为一个非常规的应用例子.况且,这个证明使德恩的定理成为一个副产品.

延长所有这些铺砖的水平边.把介于相邻的两条线之间的区域称为**条**(strip).在下面的这幅图中,有着九个正方形和五个条.

[87]

① 换句话说,当阶大于20时,矩形的完美单铺砌的种数大约是阶数的5 000 000倍.——译注



莫伦最早给出的例子

如果存在 m 个条和 n 个正方形，我们就按下述规则构造这个铺砌的 $m \times n$ **水平相交矩阵**(horizontal intersection matrix)$\boldsymbol{A}$：如果条 i 与正方形 j 的内部相交，则 a_{ij} 为 1，否则为 0. 所示这幅图的相交矩阵为

1	2	3	4	5	6	7	8	9
1	1	1	0	0	0	0	0	0
1	0	1	1	1	0	0	0	0
1	0	0	0	1	1	0	0	0
0	0	0	0	0	1	1	1	0
0	0	0	0	0	0	1	0	1

令 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ 是以那 n 块铺砖的尺寸为分量的向量，令 $\boldsymbol{y} = (y_1, y_2, \cdots, y_m)$ 是那 m 个条的高度. 令 $\mathbf{1}$ 是所有分量为 1 的 m 维向量.

定理 向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 为矩阵 $\boldsymbol{A}$ 所确定，至多相差一个常数因子，而且它们是下列方程的唯一解

$$AA^{\mathrm{T}}\boldsymbol{y} = \mathbf{1} \quad 和 \quad \boldsymbol{x} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y}. \tag{1}$$

每个正方形的高度,从而宽度,是与之相交的条的高度之和,所以由 A 的定义,$x = A^{T}y$. 同样,所有与一个给定的条相交的正方形的宽度之和一定是这个矩形的宽度 w,我们可把它取为1. 因此,仍由 A 的定义,我们一定有 $Ax = AA^{T}y = 1$.

余下来要证明的是 AA^{T} 是非奇异的,这样 x 和 y 就是唯一的了. 我们首先证明 A 的秩为 m. 要证明这一点,注意对任何两个相邻的条,不妨设为上数第 k 个条和第 $k+1$ 个条,总存在一个正方形与第一个条相交但不与第二个条相交. A 的相应列将在第 k 行有一个1,而在其随后所有的行都是0. 对所有的 k 选出这样的列(在我们前面的例子中是第2,第3,第5和第7列[①]),我们就得到了一个上三角子方阵,其对角线上都是1,因此它是非奇异的. 接下来就是一个熟知的事实,即 AA^{T} 是非奇异的. (这实际上正是我们在得出最小二乘逼近公式时用到的一个结果;例如,可参见斯特朗(Strang)的《线性代数及其应用》(*Linear Algebra and its Applications*, p.102). 这里它在一种完全不同的背景下出现,虽然这个证明也涉及square[②]!) [88]

现在可推出铺砌的有限性了. 给定了铺砖的块数,只存在有限多个可能的关联矩阵,而给定了用于铺砌的正方形的尺寸,只存在有限多种可能的铺砌.

德恩的定理也可推出. 矩阵 A 是有理数矩阵,因此(1)的唯一解一定是有理数解.

由于这个定理被建议用作线性代数的一个课程内容,所以接下来的事应该是

习题　用(1)找出莫伦的例子中那九个正方形的尺寸.

---

① 显然有误,似应为第2,第3(或第4),第5(或第1),第6(或第8)和第7(或第9)列. 另外,注意当 $k = m$ 时,所有与第 k 个条相交的正方形都符合条件. —— 译注

② square既可作"正方形"解,也可作"最小二乘逼近"中的"二乘"(即"乘方")解. 此处是双关. —— 译注

两种文化[①]

下面的内容摘自对诗人弗罗斯特[②]的一次电视采访:

> 我们从无序中脱身,升入有序,而我做的诗是一些小小的有序片段. 即便我编一个篮子,或者制一件陶器和一只花瓶,或者做什么东西…… 如果你在生活中感到痛苦迷茫,你能做的最好的事是制作一些小小的有形物,吹吹烟圈(你知道,就是那样的东西也有形状)……

[89] 对这些事,一位数学家可能会加上,"或者证明一个定理".

[90]

① 1959年,英国小说家、科学家斯诺(Charles Percy Snow,1905 - 1980)发表著名论文《两种文化与科学革命》,指出人文科学知识分子与自然科学知识分子及他们所代表的文化日趋远离,形成两种不同的文化,而这种状况将给人类带来损失. 此后,"两种文化"一词便特指这个论题.—— 译注

② 弗罗斯特(Robert Frost,1874—1963),著名美国诗人. 作品主要描写美国新英格兰地区的人物和景色. 主要诗集有《一个孩子的意愿》(*A Boy's Will*)、《波士顿以北》(*North of Boston*)等. 多次获普利策文学奖.—— 译注

第13章 蚂蚁和吉普车又回来了

蚂蚁学进修教程

吉姆·普罗普(Jim Propp)

那个勤劳的蚂蚁(见第10章)有一些姐妹和表姐妹,它们更为有趣. 这些广义的蚂蚁在一张无穷大的坐标网格图上从一个胞腔爬到一个胞腔. 在任何给定的时刻,网格图上的各个胞腔处于某种特定的状态. 我们把这些状态用0到 $n-1$ 编号,这里 n 是所允许的状态的个数. 当一个蚂蚁通过一个处于状态 k 的胞腔时,这个胞腔的状态就从 k 变为 $k+1$(以模 n 取剩余),而这个蚂蚁以它到达这个胞腔时的行进方向为基准向左转或向右转,然后离开这个胞腔. 这个蚂蚁不能自由地选择它走哪条路(向左还是向右). 它必须依照一个始终不变的长为 n 的**规则串**(rule-string)行事. 这个规则串由 n 个二进制数码组成,它们已用0到 $n-1$ 编号. 当这个蚂蚁正要离开一个原先状态为 k(而现在为 $k+1$(以模 n 取剩余))的胞腔时,如果 r_k 为 **1**,它就向右转,而如果 r_k 为 **0**,它就向左转. 这里 r_k 是那个规则串的第 k 个数码. 蚂蚁的这种广义化看来是由特克(Greg Turk)[1]最早予以考虑的,并且由布尼莫维奇和特罗别茨科伊[2]独立地予以考虑,而他们又是以科恩[3]早先的工作为基础的.

[91]

我们不妨可以把注意力集中于以 **1** 开头的规则串,因为一个以 **0** 开头的规则串可以通过把其中的二进制数码代之以其补

码的方式变为一个以 1 开头的规则串，这只不过是把“左”和“右”互换了一下，从而给出了一个镜象宇宙，这对孪生姐妹没有本质上的差别．我们将把一个以 1 开头的规则串解释成一个自然数的二进制表示．例如，原先兰顿的那个规则将被称为规则 **10**，或规则 2．容易看出，规则 **1** 是平凡的，它使一个蚂蚁绕着一个 2×2 的正方形永无休止地兜圈子．同样的道理，规则 **1**，**11**，**111**，以及如此等等，也是平凡的．

这些广义蚂蚁构成了特克等人研究的“图螨”(tur-mite)[1]的一个特例，杜德尼(A.K.Dewdney)在他的文章《二维的图灵机和图螨在一个平面上匆匆而去》(Two-dimensional Turing machines and tur-mites make tracks in a plane)中描述了这种图螨(参见[1])．这是一个很广的概念，它把图灵机也作为一个特例包括在内．结果，关于图螨的行为几乎证不出什么一般性的定理．然而，对于蚂蚁，我们至少可以证明一个结果：

定理　假定一个蚂蚁所遵循的规则串中至少包含一个 **0** 和一个 **1**，那么它的轨道总是无界的．

这个证明与第 10 章中给出的相同．

关于广义蚂蚁，我们可以提出许多问题．我将在这里讨论的一个问题是：在一个所有胞腔最初都处于状态 0 的宇宙中启动一个蚂蚁，将会发生什么情况？我们假定蚂蚁一开始头朝南．

图 13.1 表明了一个遵循规则串 **10** 的蚂蚁在一个所有胞腔最初都处于状态 0 的宇宙中徘徊了 11 000 步后的情形．状态 0 用白色画出，状态 1 用黑色画出．我们可看到西北方向上一条公路正在形成．(这条公路与第 10 章图 10.5 所示的那条公路结构相同，只是旋转了 90°．同样的图出现在杜德尼的文章中．)

---

[1] tur-mite 中的 tur 从 Turing(图灵)而来，故沿译为“图”．图灵(A.Turing，1912—1954)，著名英国数学家、逻辑学家和理论计算机科学家．关于下文提到的图灵机，可参见本译丛中《20 世纪数学的五大指导理论》中的第 4 章．——译注

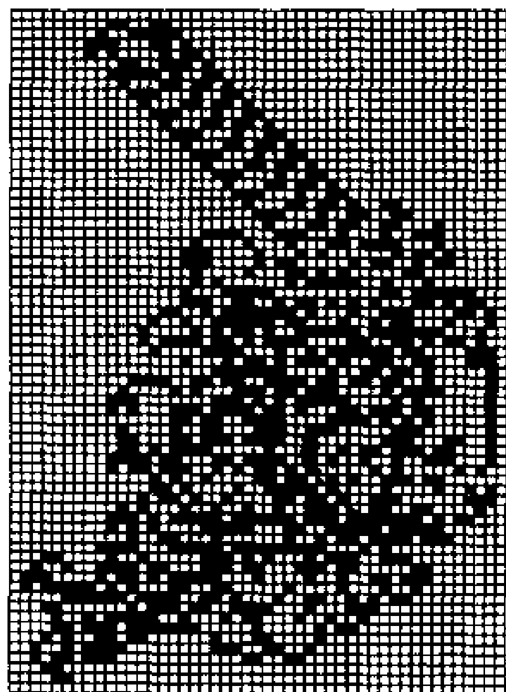


图 13.1

[92]

为了对付所循规则串的长度 $n > 2$ 的蚂蚁，我们用白色代表状态0，用黑色代表状态 $n-1$，而用介于黑白之间的不同深浅的灰色代表那些中间的状态.

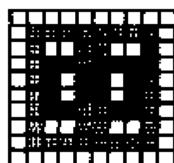
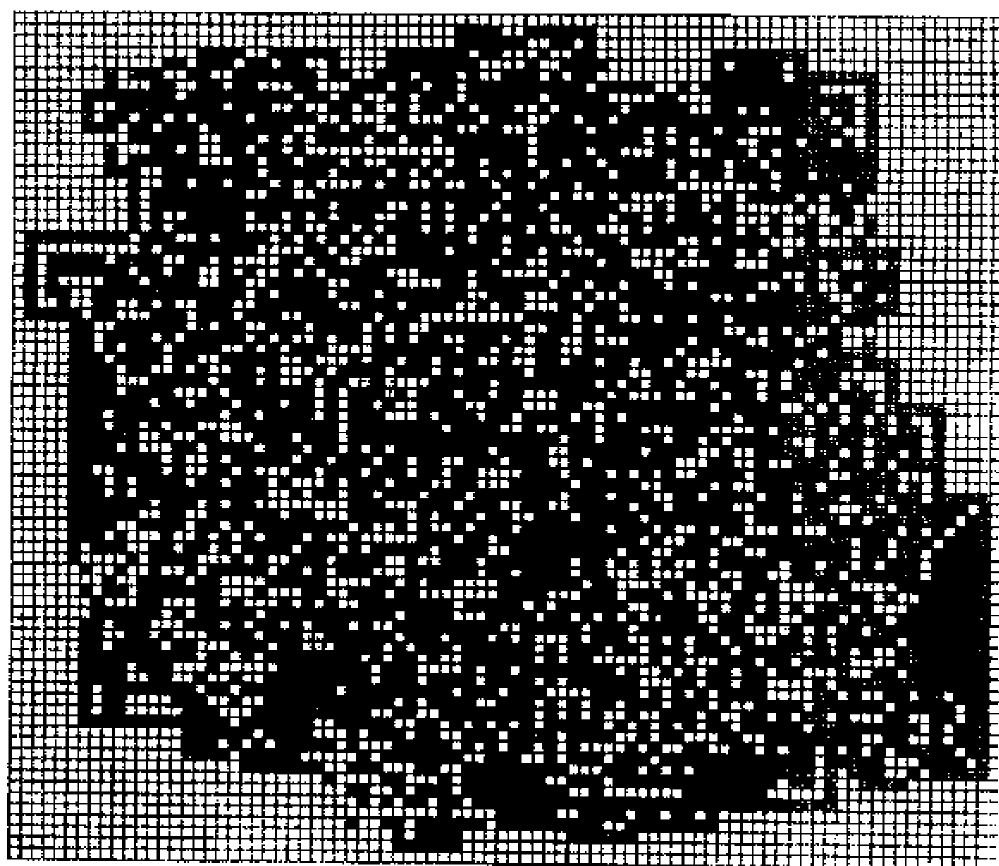


图 13.2

同蚂蚁2[①]一样，蚂蚁4在开始时创作出各种各样的对称图案(就像图13.2所示的那个，它在第236步时生成)；这些图案具有双侧对称性，这与蚂蚁2所创作的图案不同，后者是180°旋转

① 即遵循规则2或称规则10的蚂蚁. 以下类似.——译注

对称. 后来这个蚂蚁停止了对称的行为,并出了一大堆看起来乱七八糟的东西,如图 13.3 所示(100 000 步). 是不是最终会形成某种类型的公路? 我不知道. 我对它跟踪了 150 000 000 步以上,却没看到任何清晰的模式.



[93] 图 13.3

蚂蚁 5 甚至更像蚂蚁 2,因为它偏爱双重旋转对称;它的最辉煌成就是如图 13.4 所示的图案,这是走了 616 步后创作的. 然而,此后这个图案就被破坏了,而过了 150 000 000 步,人们看到了一个毫无结构重现迹象的构形. 不过,倒出现了一些令人意外的统计模式. 例如,我注意观察构形中央的一个 21 × 21 的正方形,结果发现只有 79 个胞腔处于状态 0,而有 190 个处于状态 1,172 个处于状态 2. 是什么原因导致了这种涨落?

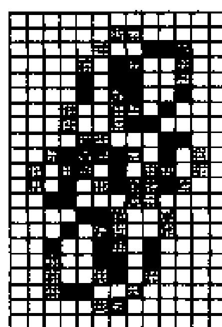


图 13.4

与蚂蚁 4 和蚂蚁 5 相反，蚂蚁 6 显得非常温顺——甚至比蚂蚁 2 还要温顺．仅仅过了 150 步，人们就能看到一条公路正在向西南伸去(图 13.5)．与蚂蚁 2 生成的那条"周期"为 104(也就是说，每建造公路的一个接续段，都是花掉蚂蚁的 104 个时间步[①])的公路不同，蚂蚁 6 形成的这条公路周期只有 18．而且实验显示，即使对宇宙的初始状态稍作一些变动，在一片状态 0 中点缀几个状态 1 和状态 2，公路还是会极其迅速地形成．

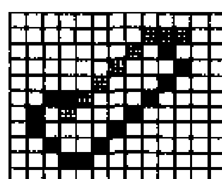


图 13.5

从蚂蚁 8 到蚂蚁 14 展现了一些新的现象．蚂蚁 8 和蚂蚁 14 完全是"混沌的"；它们建立的图案就像蚂蚁 4 和蚂蚁 5 所建立的那样，不见任何大范围结构的踪影，但它们都有别具一格的局部装饰，特别是在这朵正在不断扩展的混沌之云的边缘处．蚂蚁 10(其规则串为 **1010**) 只不过是蚂蚁 2(其规则串为 **10**) 的一个精

① 时间步(time step)，计算数学术语，即离散的时间段，这里指蚂蚁走一步的时间．注意"周期"本是一个关于时间的概念．——译注

心制作的复制品;更一般地,一个规则串如果是由一个较短的规则串通过两次或更多次的重复而构成的,那么它导致的行为将与这短规则串所导致的相同. 蚂蚁 13 开始时处于混沌状态,但是大约过了250 000步后,它便着手建造一条周期为388的公路. 蚂蚁 14 是蚂蚁 2 和蚂蚁 6 的一个奇特的杂交种;像蚂蚁 2 那样,它建造了一条周期为 52 的公路,但是这条公路看上去十分像蚂蚁 6 建造的那条. 难道这两种相似性(一种是数值上的,一种是外形上的)仅是巧合?

蚂蚁 9 和蚂蚁 12 是真正令人惊奇的东西. 在这两种情况中,图案虽然越变越大,但总是不会太偏离双侧对称模式! 更为特别的是,人们发现蚂蚁频繁地造访它出发时的那个胞腔,而当这种情况发生时,整个构形经常在蚂蚁到达那个出发胞腔的瞬间呈现双侧对称. 这种现象是由特克首先注意到的. 图 13.6 显示了蚂蚁 12 走了 16 464 步后的情况. (这幅麻省理工学院吉祥物海狸的肖象画,谨供该校教师使用.) 图 13.7 显示了同一蚂蚁走了 186 848 步后的情况,而图 13.8 显示了蚂蚁 9 走了 38 836 步后的情况. (关于这个构形在后来某一个阶段的图象,可参见[1] 的第 182 页.) 特克告诉我,吕姆勒(Bernd Rümmler) 证明了蚂蚁 12 始终在构建越来越大的双侧对称图案(参见第 18 章).

[94]

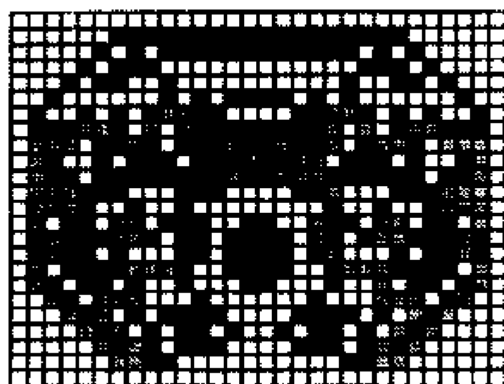


图 13.6

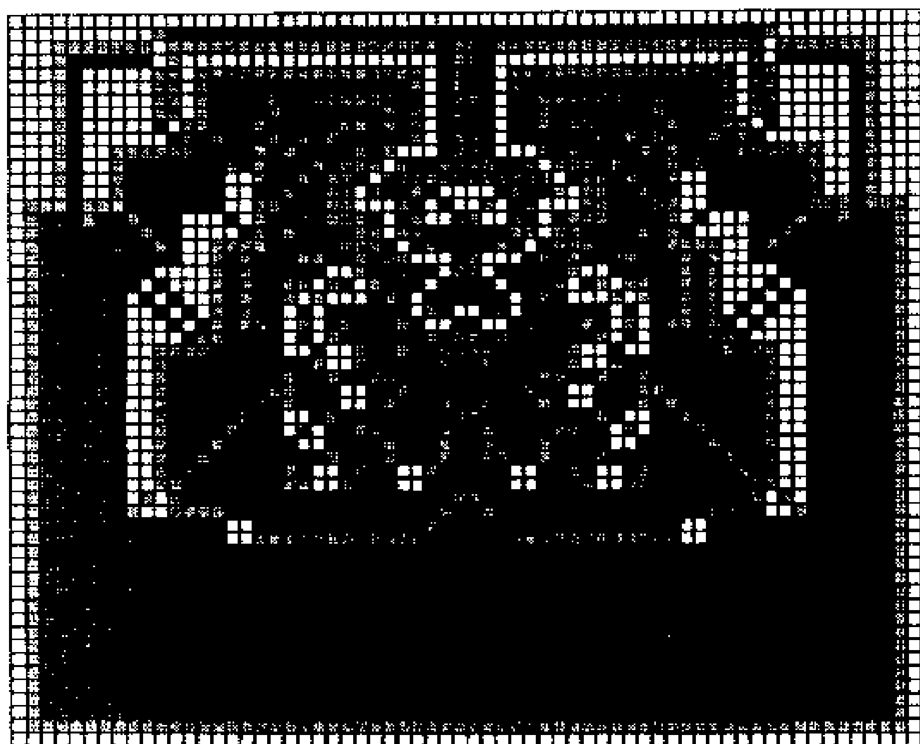


图 13.7

95

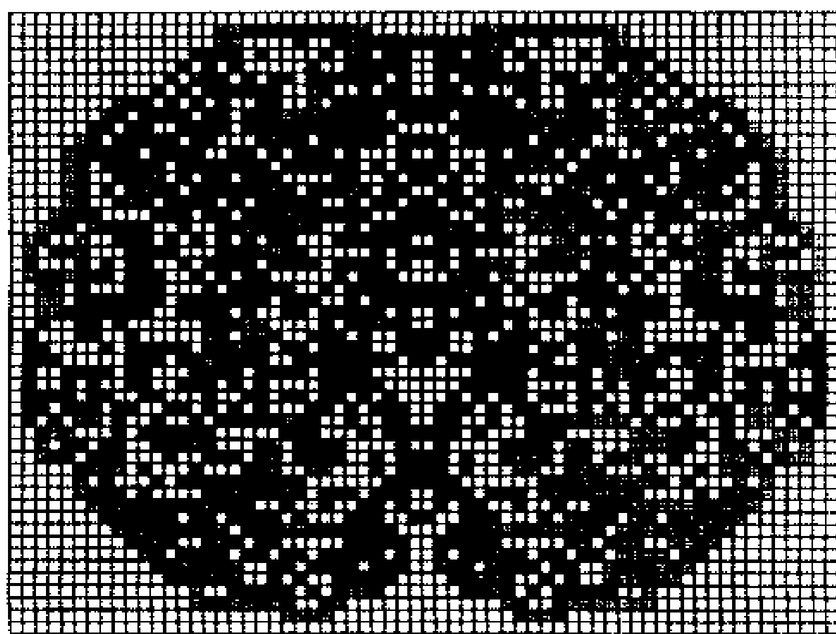
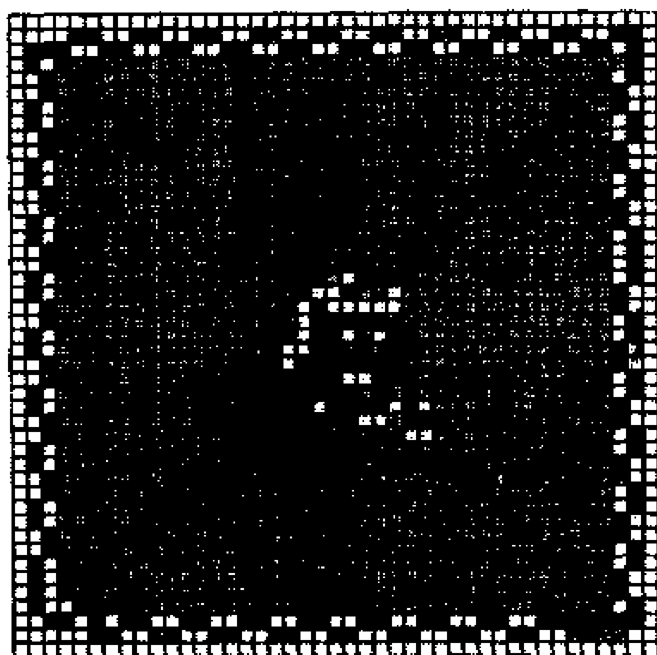


图 13.8

当我们考察有 5 个二进制数码的规则串(它对应着有 5 种状态的宇宙)时,我们所发现的新的重要行为以蚂蚁 27 为典型. 它建立了一个不断扩展着的螺旋形,如图 13.9 所示. 然而,我们却没有发现有任何蚂蚁像蚂蚁 9 和蚂蚁 12 那样在构建不断扩大的双侧对称图案. 为了再找出这种类型的蚂蚁,我们不得不走进长度为 6 的规则串. 在这里,我们遇到了另一个谜团:长度为 6 的规则串中导致双侧对称图案的有 33,39,48,51,57 和 60. 注意所有这些数都能被 3 整除! 这肯定不是一种巧合.



[96] 图 13.9

关于广义蚂蚁的研究可以很容易地变为两队人马之间的一个对策游戏. 一队是"理论家",他们将设法揭示一般的规律,这些规律将描述什么样的长期行为是可能的,以及给出一个特定的蚂蚁规则和一个特定的宇宙初始状态,将会发生哪一种行为. 另一队是"工程师",他们将设法在蚂蚁的宇宙中设计图案,这些图案可用来作为一种通用计算机的建筑模块. 如果工程师们取

得了成功,那么就可以证明理论家们的目标在某种意义上是不可达到的,就像在康韦的生命游戏中的情况(参见[4]的第25章). 甚至可能是这种情况:像"蚂蚁4究竟会不会建造一条公路"这样简单的问题,在你那爱不释手的数学公理基础(ZFC[①]或诸如此类的公理系统)中也是不可证明的.

希望读者自己对这些蚂蚁多加玩味并得出自己的结论.

在此谨向布尼莫维奇、科恩、X. P. Kong、兰顿、布鲁斯·史密斯(Bruce Smith)、特罗别茨科伊、特克和 Fei Wang 表示感谢,本章引用了他们的工作.

参考文献

1. A. K. Dewdney, Computer recreations, *Scientific American* September (1989), 180—183; follow-up March (1990), 121.
2. L. A. Bunimovich and S. E. Troubetzkoy, Rotators, periodicity, and absence of diffusion in cyclic cellular automata, *Journal of Statistical Physics* 74, January (1994).
3. E. G. D. Cohen, New types of diffusion in lattice gas cellular automata, in *Microscopic Simulations of Complex Hydrodynamic Phenomena*, M. Mareschal and B. L. Holian ed., Plenum Press, 1992.
4. E. R. Berlekamp, J. H. Conway, and R. K. Guy, *Winning Ways*, Academic Press, 1982.

吉普车回来了[②]

许多读者无疑知道所谓的吉普车问题[③],但如果你们中可

① 即策梅洛-弗兰克尔(Frankel)公理系统加上选择公理所形成的公理系统. 可参见《公理集合论导引》(张锦文著,科学出版社,1991年版).——译注

② 本节未按例署名,但看来是盖尔的文章.——译注

③ 此题的一个简单形式曾取作1962年北京市中学数学竞赛试题,见于《北京市中学数学竞赛试题汇集》(北京市数学会编,科学普及出版社,1964年版).——译注

能有谁忘了的话——这问题是要让一辆吉普车穿过一片沙漠.困难在于这辆吉普车所带的燃料只够跑部分路程.解答中允许吉普车预先进入沙漠数次,以在沿途的一些储藏点存放不同量的燃料,这样它就可以在最后的运行中根据需要补充燃料.这辆吉普车能不能用这种方法穿过任意长度的沙漠?如果能,那么怎样以最小的燃料消耗量做到这点?

对这个问题的一个完整的解答由法恩(N.J.Fine)于1947年给出.自那时以来,人们讨论了这个问题的众多变种.特别是,鄙人于1970年指出,如果这问题中派遣的不是一辆而是 n 辆吉普车,那么最小的燃料用量将严格地小于单单一辆吉普车的用量的 n 倍(《美国数学月刊》同意在这里刊出这个结果,甚至允许这篇文章加上"吉普越多越俭朴"(Jeeper by the Dozen)[①]这个副题)[②].在那篇文章中,我把下面这个看来很自然的问题列为尚未解决的问题:假定在行程的两端都有燃料供应,要求以最节约的方法让一辆吉普车穿过沙漠又回到出发点.我很高兴[97]地宣布,现在,即过了23年之后,这个问题已被州立博伊西大学的豪斯拉思(Alan Hausrath)和州立圣何塞大学的杰克逊(Bradlay Jackson)、米切姆(John Mitchem)和施迈歇尔(Edward Schmeichel)所解决.他们的解答十分精美,我将在这里试图予以定性的描述.然而,在做这件事之前,我将对原来的那个吉普车问题再作一次考察,正如有人指出的那样,可以用一些经验常识上的理由使它的解答显得十分合情合理.

首先,考虑一个显然等价的问题结果将更为方便——计算一辆吉普车用它燃料箱容量的 x 倍的燃料所能走的最长距离.从现在开始我们将一直采用这种表述.

---

① jeep,一义"吉普车",音译,名词;另一义"小的",俚语,形容词.这里 jeeper 是后者的比较级,但词形上又扣前者,系双关.故权作此译.——译注

② 参见附录2.——原注

我猜想大多数人在思考这个问题时会构想这辆吉普车在出发点与各个储藏点之间来回奔走，并在沿途存放或取走燃料．然而，不妨假定每次吉普车回到大本营时都由另一辆吉普车来代替它做下一次出行．这当然不会影响这个问题的实质．更进一步，这辆新的吉普车还可以由一位新的司机驾驶．如果真是这样的话，人们就可以节省大量的时间，因为没有理由要每一辆吉普车等到前一辆车回来之后才出发．它们可以一起离开出发线，组成一个护运车队一路同行．除了一辆以外，所有吉普车的任务都是为其他车辆补充燃料．图 13.10 是一个四车护运车队的一幅简略鸟瞰图，它们从 S 出发，向 F 前进．

我们把一辆吉普车用一箱燃料能走完的距离取作距离的单位．在图中，那辆带格子图案的吉普车(或称超级吉普车) J^* 被指定要走完全程，其他都是燃料补给车．这些燃料补给车必须回到出发线，这相当于假定每单位距离它们所消耗的燃料是那辆超级吉普车的两倍．吉普车上阴影部分代表它们走过距离 x 时燃料箱中所剩的燃料．因此非阴影部分对超级吉普车来说代表了 x 单位的燃料，对其他车辆则代表了 $2x$ 单位的燃料．在图 13.10 的情况发生的时候，1 号吉普车正要把它那 $(1-2x)$ 单位的燃料全部供出去，这些燃料正好把其他三辆车的燃料箱注满．

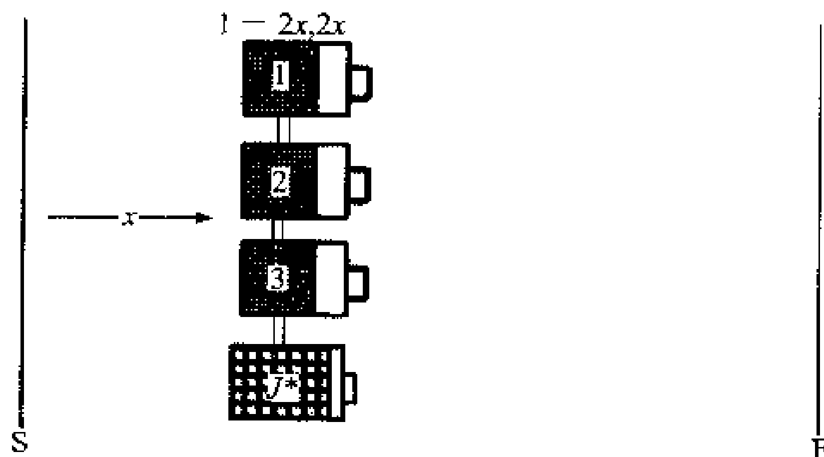


图 13.10

所以$1-2x=5x$,即$x=1/7$. 于是这个四车护运车队的问题就变成了一个三车护运车队的问题. 用这种方法,现在容易知道一个 n 车护运车队能走的距离为

$$1+\frac{1}{3}+\frac{1}{5}+\cdots+\frac{1}{2n-1}.$$

如果给出的燃料量不是一个整数,则令 f 为其小数部分,并在护运车队中另外增加一辆燃料补给车带着这额外的 f 单位燃料. 一个同上面类似的计算表明,人们可以再得到 $f/(2n+1)$ 单位的距离. 为了下面的应用,我们定义 $D(x)$ 为一辆吉普车用 x 单位的燃料所能走的最长距离. 于是有

$$D(x)=1+\frac{1}{3}+\cdots+\frac{1}{2n-1}+\frac{f}{2n+1}, \qquad (*)$$

其中 f 是 x 的小数部分,而 $n=x-f$.

吉普车必须穿过沙漠再回来的问题称为**来回旅行问题**(round-trip problem),而这里的公式还要简单,即 $1/2+1/4+\cdots+1/(2n)$,用护运车队的模型很容易把它推出. 同样,对于 k 辆(或称多辆)吉普车的问题,同一模型表明,用 n 箱燃料,其中 $n>k$,这 k 辆吉普车能走的距离为 $1+1/(k+2)+\cdots+1/(2n-k)$. 当然,在所有这些情况中,还需要某种进一步的论述来证明这些公式确实给出了最优的距离.

在进而描述我那所谓的**双向吉普车问题**(two-way jeep problem,即在两个端点都有燃料供应的来回旅行问题)的解答之前,让我暂时离开一下正题,去考虑一个更为一般的护运车队问题. 前面我们假定所有的吉普车都是一样的,现在我们考虑吉普车有各种各样的类型. 一辆吉普车的特征由两个数所刻画,一个是它的**容量** $C=$ 它所能承载的燃料的升数,另一个是它的**燃料效率** $E=$ 它用一升燃料所能走的公里数. 于是就有一个问题:给出 n 辆吉普车,每一辆都有各自的 C 和 E,能被穿越的沙漠最长是多少?

我相信有一类关于这个问题的文献(背景通常是火箭和星际空间,而不是吉普车和沙漠),不过据我所知,一般性的问题尚未解决,这意思是说,还没有任何已知的“好算法”来找出最优解. 可能会存在某种燃料补给系列来实现这个最优解,然而那是什么呢? 把这种情形予以形象化的一个方法仍由图 13.10 给出,但这次请把吉普车想像成火箭,而且这幅图处于一个垂直的平面而不是一个水平的平面. 这样燃料补给就可以连续地进行. 重力使燃料从最上面那支火箭中源源流出,以让下面各支火箭始终保持燃料满载状态. 当最上面火箭的燃料箱空了的时候,这支最上面的火箭就被抛弃,而其他的火箭则继续前进,并重复这个过程.

可以想像这个问题是 NP – 困难的[①]. 就像许多最优化问题那样,困难在于不存在一种容易的方法来识别什么时候一个给出的解才是最优的. 人们怎样来对付这种情形呢? 有一个一般的方法,有时会起作用. 人们去证明任何最优解必定满足某些条件(例如,一个可微函数在其内部极大点上导数必定为零). 如果运气好的话,人们会找到足够多的这类条件,以致有一个唯一的对象满足这些条件,因此这个对象就一定是最优解. 事实上,这正是那些作者对双向吉普车问题所做的事. 我下面即予描述.

我们首先考虑当 S 处有 f 单位的燃料而 F 处有 g 单位的燃料时能穿行的沙漠最长是多少. 下面的推理多少带着些常识性的判断. 如果 $g \geqslant f$, 那么人们能得到的最好结果显然是 $D(f)$(由(*)式给出),这时人们把 S 处的燃料用于前行,而把 F [99]

① NP – 困难的问题是指一类可由任何一个 NP 问题用多项式时间转化而来的计算搜索问题. 因此,NP – 困难的问题至少同 NP 问题一样难;如果一个 NP – 困难的问题同时又是一个 NP 问题,那它就是一个 NP 完全问题. 关于 NP 问题和 NP 完全问题的一个通俗而又不失准确的介绍,可参见译者在《科学》2001 年第 4 期上的文章《扫雷高手的百万大奖之梦》. —— 译注

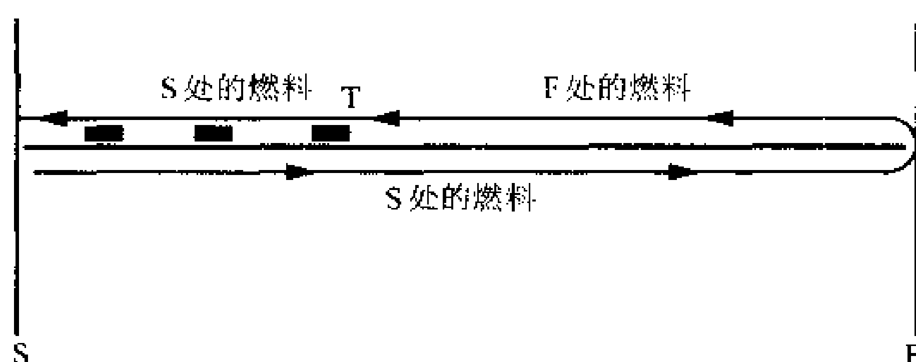


图 13.11

处的燃料用于返回．另一方面，如果 $g < f$，这个过程显然不是最优的，因为这样人们能够得到的仅是 $D(g)$ 而已，而 S 处的燃料却还有一些没有用到．因此，一定有某种方法在前行的途中建立一些储藏点以存放 S 处的燃料，这些燃料可以在归途中使用；所以我们设想这样的储藏点已经建立，而且假设其中离 S 最远的那个在点 T，如图 13.11 所示．

这些作者接着便证明存在着具有下述性质的最优解，而且这些性质看起来十分合情合理．很显然，在前行的过程中只能用 S 处的燃料，但他们证明存在一个最优解，它满足（ⅰ）在归途中从 F 回到 T 只用了 F 处的燃料，（ⅱ）从 T 回到 S 只用了 S 处的燃料（如图 13.11 中所标示的）．从直觉上看，把 S 处的燃料带到 F 处（这相当于把煤带到纽卡斯尔[1]），或者把 F 处的燃料带回到比 T 离出发处更近的地方（如果你需要这些额外的燃料，你本该在前行的途中把它们存在那儿），显然都很浪费．

然而，结果是这些条件实际上决定了这个解．具体地说，从 T 到 F 的距离 d_1 必定是 $D(g)$，因为根据（ⅰ），在从 F 走到 T 的过程中必须把全部的 g 都用掉，而且除此之外什么都不用．因

[1] 纽卡斯尔，即泰恩河畔纽卡斯尔，英格兰东北部海港城市，曾主要出口煤炭．此句原文为 coals to Newcastle，系习语，也可译成其引申义“多此一举”．——译注

此，在前行时吉普车必须把 g 单位的燃料运送到 T. 从 S 到 T 的距离 d_2 必定是一辆吉普车被要求用 f 单位的燃料把其中的 g 单位运到目的地后再回来时所能走的最优距离. 把原来那个来回旅行问题稍作修改，就变成求这个最优距离的问题，用护运车队的方法很容易把它解决. 于是 $d = d_1 + d_2$ 就是我们所要寻求的距离. 这样，要解决这个双向问题，只要把原先那两个吉普车问题的解拼合在一起即可.

这些作者接下来又讨论另一个双向问题：给出 x 单位的燃料，要走出最远的距离，人们应该怎样选择 f 和 g？其中 $f + g = x$. 对于 $2 \geqslant x$，人们能做的最好的事是每个端点各放一半燃料. 对于 $x > 2$，一般来说解并不是唯一的，但总存在这样一个解，其中 g 是由 $g = [((x+1)/2)^{1/2}]$ 给出的一个整数. 当 x 的值较大时，这个 g 比燃料量的一半小得多，尽管所走的最优距离与每个端点各放一半燃料时所走的距离相差不到 $1 + \ln 2$（与 x 无关）. 而且，归途中所需要的途中储藏点的个数是 $[x] - 2$. 这样，图 13.11 就代表了燃料供应量介于 4 箱和 5 箱之间时的解①.

当然，证明这些作者的算法的最优性是他们文章的主要内容，而我们在这里所做的只是描述这个算法是怎么回事. [100]

① 此处疑有误. 当 $4 < x < 5$ 时，$[x] = 4$，$[x] - 2 = 2$. 但图 13.11 中却有 3 个储藏点. 事实上，用护运车队方法可知，若一辆吉普车用 f 单位燃料把其中 g（整数）单位送到最远的终点再回来，则当 f 是整数时，途中储藏点加上终点（即图 13.11 中的 T 点）共 $f - g$ 个，当 f 不是整数时，共 $[f] - g + 1$ 个. 而用 g 单位燃料从 T 走到 F 或从 F 走到 T，途中各设储藏点 $g - 1$ 个. 从 T 到 F 的储藏点与从 F 到 T 的不同，故从 T 到 F 再回到 T，储藏点一共 $2g - 2$ 个. 这样，整段路程的储藏点总数，当 f 是整数时，为 $(f - g) + (2g - 2) = x - 2$，当 f 不是整数时，为 $[x] - 1$. 这两种情况可统一表示为 $\lceil x \rceil - 2$. —— 译注

第14章 围　棋

问　题

本书第11章专门讲了某几种游戏，而且我们指出，虽然对策论已经成为纯粹数学和应用数学的一个庞大的分支学科，但是除了极少数几个例外，数学文献中所研究的对策游戏的集合与人们实际上玩的游戏的集合并不相交．一个著名的例外就是佩特森和兹维克对儿童游戏"记忆翻牌"的分析，这两位作者实际上发现了其中最优玩法的规律；也就是说，他们解决了这个对策问题．

本章也是专讲一个人们平时玩的游戏，事实上，这是在人类历史上最常玩和最重要的游戏之一，它就是围棋．我们的报告基于伯利坎普及其合作者沃尔夫（David Wolf）、默夫斯（David Moews）、金龙焕（Yonghoan Kim）和 Raymond Chen 最近的某些一连串的工作．然而，我们应该马上指出，与佩特森和兹维克的工作不同，伯利坎普并没有解决围棋对策问题．事实上，他的工作对人们乃至计算机在改进游戏策略方面即便提供了什么帮助，那么到底有多少帮助也是不清楚的．他所做的是发明了一系列围棋排局问题，这些问题计算机程序能够解决，但是还没有一位职业棋手能够成功地予以解决．图14.1中摆出的盘面就是一[101]个例子．

就其本身来看，这个盘面可能没有什么令人感到特别的地

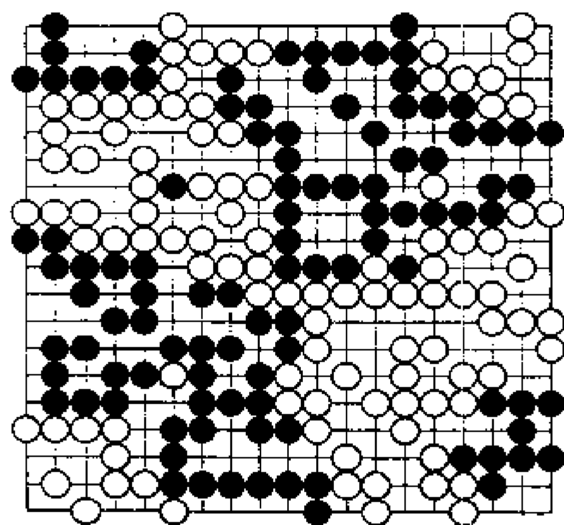


图 14.1 白方先行且必胜

方. 其实,有一些在计算机的帮助下发明的国际象棋排局的例子,人们甚至不期望卡斯帕罗夫[①]有能力予以解决. 图 14.2 给出了一个例子.

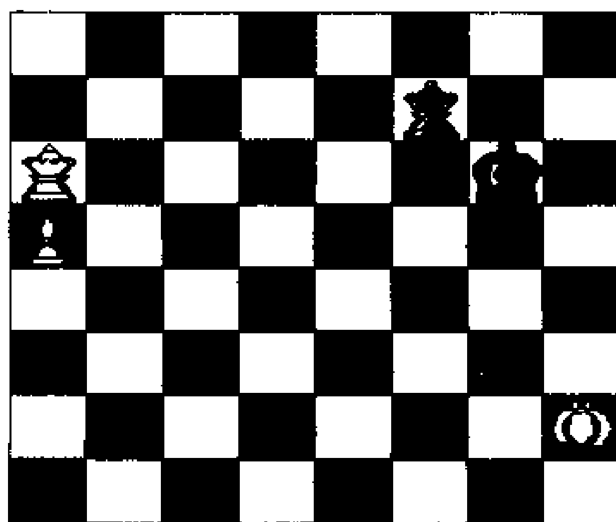


图 14.2 黑方先行,白方在 109 步内必胜(通过兵升变,吃子或将杀)

[102]

① 加里·卡斯帕罗夫(Гарри Каспаров,1963—),世界著名的俄罗斯国际象棋国际特级大师,常年保持世界冠军称号.——译注

有哪一位人类棋手会认识到这是一个对白方来说必胜的局面,这一点看来令人怀疑,因为黑方能走那么多步来负隅顽抗.然而,这个排局问题除了是一个稀奇之物,几乎没有什么令人感兴趣的地方.与此相反,使伯利坎普的工作备受瞻目的是,为解决他的问题,人们必须应用一个广泛、深刻而美丽的数学理论,这个理论被它的创建者伯利坎普、康韦和盖伊称为"组合对策论"(combinatorial game theory).世界上只有少数几个数学家知道这个理论.它在这些作者那850页的两大卷著作《取胜之道》中有详细的阐述.在介绍它对围棋的应用之前,我将试图给出这个理论是怎么回事的一个简短概述,也就是说,把850页的内容总结在少数几段文字中——这显然是不可能做到的,但是无论怎样,这里做了.(平心而论,应该说那本书中只有约150页是专讲这个理论的,大部分篇幅是关于它对具体游戏的应用.)

组合对策论

这个理论研究的只是非赢即输的游戏,这种游戏在左先生和右太太之间进行,走最后一步的人就是赢家.虽然这看起来有点具体,但许多游戏只要在规则上作些简单的变化就可以成为这种形式.这样的游戏可以十分方便地用树来表示,就像图14.3中的那棵.

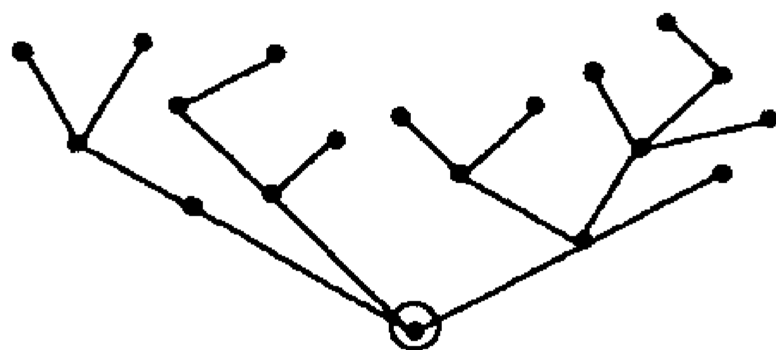


图 14.3

如图所示，那颗“纽扣”最初放在树根上．若轮到左先生走，这颗纽扣就沿着左边的某条边被推前一步；若轮到右太太走，这颗纽扣就沿着右边的某条边被推前一步．把它推到一个终端顶点的人就是赢家．

这种游戏的一个简单而绝非平凡的例子，是“横行霸道”（domineering），它可用图 14.4 来例示．

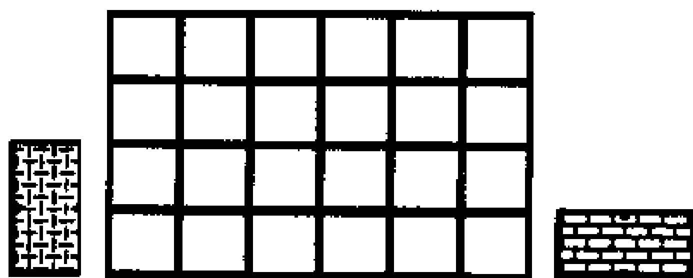


图 14.4

规则十分简单．左先生手中有着垂直的多米诺骨牌，右太太有着水平的多米诺骨牌；他们轮流把骨牌放到那方格图上——一个垂直地放，一个水平地放，让每一张骨牌正好覆盖两个方格．骨牌不允许重叠．如果一位局中人不能再放了，他就输了——就这些．我选择 4×6 的“横行霸道”作为一个例子，是因为在我写这篇文章的时候，这个游戏还没解决．

回到一般的理论．所有的游戏都属于 4 种类型之一：

（1）一个对左先生来说是必胜的游戏（不管谁先走）；

（2）一个对右太太来说是必胜的游戏（不管谁先走）；

（3）一个对第一位局中人来说是必胜的游戏；

（4）一个对第二位局中人来说是必胜的游戏．

这 4 种可能性的每一种都可用图 14.5 中的“横行霸道”游戏来例示．

我们把图 14.5（d）看作一个第二位局中人必胜的游戏，是因为第一位局中人在他开局走第一步时就不能放了．关于第二

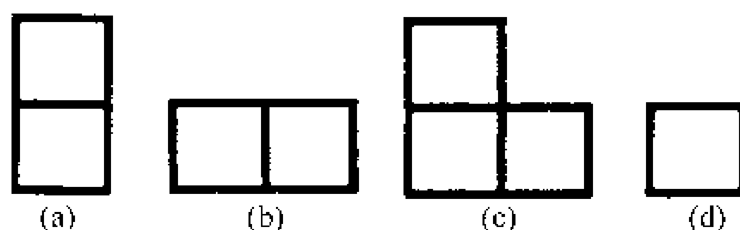


图 14.5

位局中人必胜的游戏的一个稍稍不平凡的例子是图 14.6 所示的棋面,图中还画出了相应的树[①]. 读者很容易验证这是一个第二位局中人必胜的游戏.

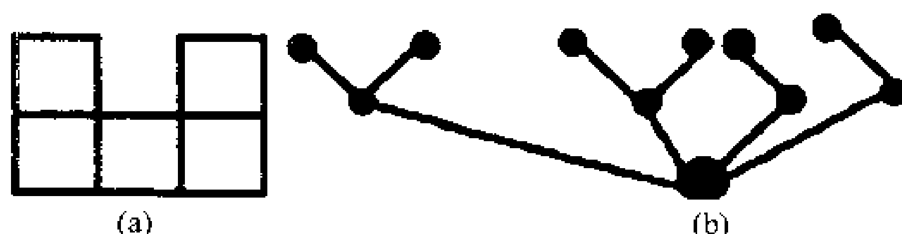


图 14.6

首先的一个考察结果是,对所有游戏的集合,有着一个自然的**加法**概念. 具体地说,就是把两个游戏 G 与 H 之和定义为这两个游戏同时玩. 以“西洋跳棋”+“连城”为例(须作适当修改以使它们成为非赢即输的游戏). 比方说,左先生执黑子,画×;右太太执红子,画○. 轮到一位局中人走时,他可以在其中任一个游戏中走一步.

其次,把一个游戏 G 的逆 – G 定义为局中人地位被互换了

① 注意从树的任一个顶点出发的边中,左边是只能由左先生走的边,右边是只能由右太太走的边. 左先生在开局走了一步后,即他沿左边那两条从树根出发的边中的一条走了一步后,若允许他再走一步,则由图 14.6 左的方格图可知,他还是能走的,因此从这两条边的顶点出发都有一条向左的边(当然,还各有一条向右的边让右太太走). 而右太太在开局走了一步后,即使允许她再走一步,她也不能走了,因此从右边那两条边的顶点出发就没有向右的边了.——译注

的 G(相当于以一条垂直轴为对称轴把树作镜射).

接下来是关键的考察结果:我们定义 G[①]上的一个等价关系 ~[②]. 如果 $G-H$[③]是一个第二位局中人必胜的游戏,我们便写作 $G \sim H$. 注意 ~ 是自反的:第二位局中人可用"依样画葫芦"的方法赢得 $G-G$——第一位局中人在 $G(-G)$中怎样走,他在 $-G(G)$中也怎样走. 从现在开始,我们将用 $G \sim 0$ 来代替 G 是一个第二位局中人必胜的游戏的说法. 现在必须证明下面的:

(1) 如果 $G \sim 0, H \sim 0$,则 $G+H \sim 0$(你看出了这一点吗?)

(2) 如果 $G \sim 0$,则对所有的 $H, G+H \sim 0$. 这是因为根据自反性和(1),$(G+H)-H=G+(H-H) \sim 0$.

(3) 对称性:当且仅当 $H \sim G$ 时,$G \sim H$(因为 $-(G-H) \sim H-G$,等等).

(4) 传递性:如果 $G-H \sim 0, H-K \sim 0$,则 $G-H+H-K \sim 0$,所以 $G \sim K$.

结论(你看到它是怎么来的)是等价类的集合形成了一个阿贝尔群,它的零元素是所有第二位局中人必胜的游戏的集合. 我们把这个群称为 Γ,这就是组合对策论的研究对象. 从现在开始,G 代表一个等价类,因此 ~ 可用 = 代替. 为了获得对 Γ 的一个感觉,我们考虑图 14.7 中的"横行霸道"游戏.

这两个游戏显然都是左先生必胜的游戏,但图 14.7(a)是与图 14.5(a)相同的游戏(请证明这一点),而图 14.7(b)则不是. 图 14.5(a) - 图 14.7(b)是一个左先生必胜的游戏,而不是第二位局中人必胜的游戏. (如果你没有成功地验证这些结论,你就失去了一半的乐趣.)

---

① 符号运用混淆. 这里的 G 是指所有游戏的集合. ——译注

② 关于等价关系以及下面说到的群、阿贝尔群(即交换群)、子群、偏序等概念,可参见《现代数学选讲》(李绍宽著,中国纺织大学出版社,2000年版). ——译注

③ 即 $G+(-H)$. ——译注

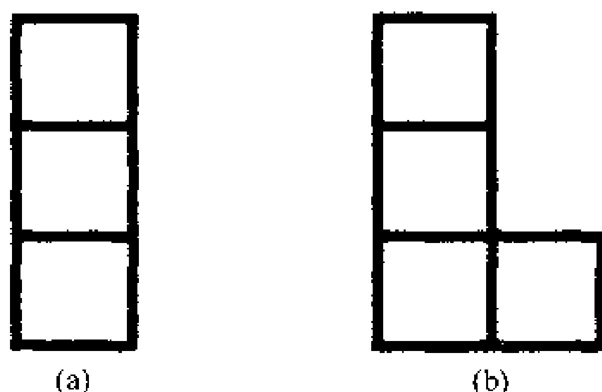


图 14.7

回到这个理论．在 Γ 上有一个自然的偏序．如果 G 是一个左先生必胜的游戏，我们就称 G 是正的，即 $G>0$；如果 G 是一个右太太必胜的游戏，则称 G 是负的．注意 > 是良好定义的，也就是说，等价于一个正游戏的游戏是正的（为什么？）．显然，正游戏之和是正的（左先生应当总是在右太太走子的那个游戏中应对），因此定义若 $G-H>0$ 则 $G>H$ 就给出了 Γ 的一个偏序．这样，每一个游戏或是正的（图 14.5(a)），或是负的（图 14.5(b)），或是零（图 14.5(d)），或上面这三种情况都不是，这意味着它是一个第一位局中人必胜的游戏（图 14.5(c)）．[105]

好了，既然我们有了一个偏序阿贝尔群，我们打算用它干什么呢？到这一步，有关的分析发生了一个令人意外的转折．在通常研究群论的代数方法中，人们将着手研究这个群的结构、它的子群及表示，等等．但这里的兴趣不在结构而在单个的群元素．事实上，这些元素有点像一部戏剧或小说中的人物那样组成一副场景登场了．它们甚至有名字，有的叫"数"，有的叫"星"、"上"、"下"、"正小不点儿"（Tiny）和"负小不点儿"（Miny）（符号分别是 $*$，$\uparrow$，$\downarrow$，$+_{on}$，$-_{on}$），此外还有十几个名字．更有甚者，这些元素的每一个不但有一个名字和一个符号，而且还有一幅"图"，即它的树．当然，许多不同的树对应着同一个游戏．例如，图 14.6(b)中的树代表游戏 0，而它的树只是一个单独的

顶点．这个理论的一个重要定理说，每一个游戏都有一个唯一的典范型，这意思是说，唯一的一棵具有尽可能最少的分支的树．例如，图 14.7(b)的典范型由图 14.8 给出．

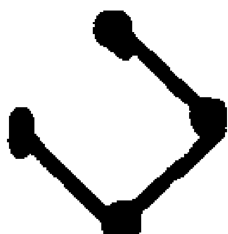


图 14.8

Γ 具有许多有趣的子群．最简单的或许是{0, *}，其中 * 就是图 14.5(c)中的游戏，它的树在图 14.9 中给出．

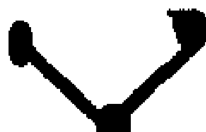
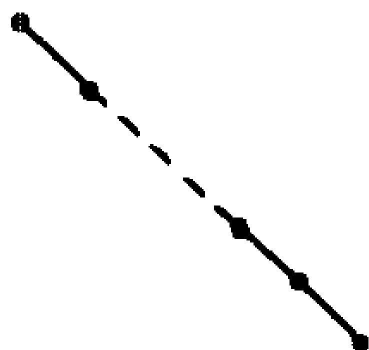


图 14.9

你应该验证 $* \dot{+} * = 0$．事实上，树呈左右对称的任何游戏显然必定具有阶 2．图 14.3 中的树我只是画来玩玩的，结果发现它的典范型就是 * 的树(图 14.9)．这是伯利坎普给我指出的．一个显然的问题是：给出一棵树，怎样找出它的典范型？这个一般的问题是 NP－困难的．

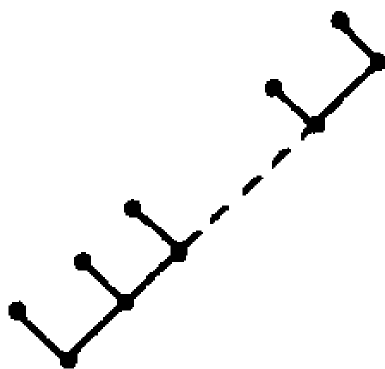
Γ 的最重要的子群是一个称作"数"的群，它同构于二进有理数群[①]．n 的树只不过是

① 无穷级数 $a_{-m}2^{-m} + a_{-m+1}2^{-m+1} + \cdots + a_{-1}2^{-1} + a_0 + a_1 2 + a_2 2^2 + \cdots$(其中 a_i 为 0 或 1，m 为任一正整数)在某种"绝对值"下将"收敛"到一些实数，此即二进有理数，如所有整数、$1/2^n$，3/4 等．它们在加法下组成的群即二进有理数群．详情可参见《p 进数》(冯克勤著，湖南教育出版社，1995 年版)．——译注

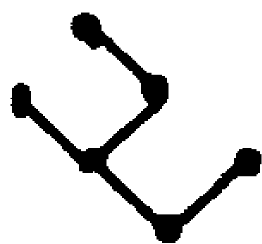


[106]

而证明 $m + n = (m + n)$ [1] 是一件微不足道的事. 更有趣的是 $1/2^n$. 它的树是



于是,图 14.8 和图 14.7(b)都是 1/2. 一个美妙的练习是证明 $1/2^n + 1/2^n = 1/2^{n-1}$. 因此,"数"的加法就像普通数的加法一样. 3/4 的树是



请回忆 Γ 是被赋予偏序的. 因此人们可能要问,比方说, * 在这个偏序中地位如何. 可以证明 * 是一个"无穷小",这意思

① 这里 m 是指由正整数 m 代表的游戏(等价类,下同), n 是指由正整数 n 代表的游戏,而 $(m + n)$ 是指由 $m + n$ 代表的游戏. —— 译注

是说,它小于任何正"数"而大于任何负"数",但与0不可比较.这是一个很容易的练习.

现在我的问题是决定何时打住,因为还有如此多的内容可以介绍.例如,游戏具有叫做**温度**(每一个游戏或者是热的,或者是冷的,或者是温的)和**原子量**的性质.从Γ到"数"有着重要的同态,它们叫做**凉快**(cooling)和**激冷**(chilling).在一些情况下,**激冷**有一个逆射,叫做**暖和**(warming);等等,等等.不过,现在我们该去说明与围棋有关的内容了.

解决与分解

一个游戏被解决了,这意思是指什么呢?一个对组合对策论(以下简称为CGT)还没入门的人可能会以为,这是指知道谁是必胜者并且知道怎样才能取胜,也就是说,描述出**取胜之道**.CGT对"解决"的观念颇有些独特,而且首先就与这种幼稚的观念多少有点不同.说一个游戏被**解决**了,意思是指人们确切地知道了这个游戏是Γ的哪一个元素,或者更准确地说,知道了它的典范型.这被称为这个游戏的**值**.(在实际操作中,如果能把一个游戏表示为一些其值已知的游戏的和,这将是令人满意的.过一会儿我们还要提到这一点.)

107

举一些例子:2×6"横行霸道"的值是$1-\frac{1}{2}-1$(如果你不知道这符号的意思,请不要感到不安).已知5×5"横行霸道"具有值0,因此它是一个第二位局中人必胜的游戏,然而,知道了这件事实并不意味着人们不用执行一个大运算量的计算就会得到这第二位局中人的必胜策略.

事实上,关于"横行霸道"的情况颇为令人迷惑.甚至像2×n"横行霸道"这样看上去十分简单的游戏也是很不平凡的.伯利坎普对当n为奇数时的值有一个"公式".结果发现只有当n=13时2×n才是一个第二位局中人必胜的游戏.在撰写《取胜之道》的时候,4×4尚未解决,而在前一年,伯利坎普曾把它作

为一项学期作业布置给麻省理工学院的两名大学生．他们给出了不同的答案，没有一个是正确的．我向埃尔温[1]讨教正确的解答，他提供了，但是结果发现这解答还是有个毛病．最后，沃尔夫利用一个计算机程序给出了正确的（但愿如此）典范型．这棵树竟有 52 个分支．

但如果 CGT 不能告诉人们怎样玩游戏，那它有什么好？这是一个合理的问题，回答是这个理论对一类受到很大约束的游戏极其有用，具体地说，它们是 Γ 中那些能表示为其值已知的简单游戏之和的游戏．这种游戏称为**分解的**．它们的典型就是 Nim 游戏，这里人们显然是在玩一些游戏的一个和，而其中每一个游戏都是彻底平凡的．（Γ 中对应于一个 n 子 Nim 堆的元素称为 Nimber，记为 $*n$．习题：画出三子 Nim 堆的树．）其实，一种看待 CGT 的方法是把它看作 Nim 理论的一个大范围推广．

现在我要说，大多数可能使人们入迷的游戏都不是可分解的．例如，第 11 章中所描述的“大嘴巴”游戏就决不会被分解．然而，谁玩过那著名的“连点成盒”游戏，谁就曾遇到过关于分解的一个绝妙的例子．请回忆一下在残局阶段，盘面被分割成不相交的区域，它们具有这样的性质：在一个区域的两点间连上一条线，就把这个区域中的所有盒子都给了对方．（在《取胜之道》中，有整整一章共 43 页专门用来讨论“连点成盒”，这是伯利坎普的一个特产，虽然还没有建立起一个完整的理论．）

那么（最后！）关于围棋的情况如何？在一局围棋开始阶段出现的局面肯定不能分解．然而，在终盘的最后阶段所出现的局面确实可以分解——既在 CGT 的专门性意义下，也在盘面被划分为各自独立的区域这个地理意义下——于是应用 CGT 就能得到丰硕的成果．对于仅包含少数几个“地块”（区域）的局面，即使其中有些“地块”仍超出了 CGT 分析力所能及的范围，

---

① Elwyn，伯利坎普的名．如此称呼表示关系密切．——译注

职业围棋手还是能十分娴熟地予以把握．伯利坎普和他的合作者所做的是构造了某些围棋局面，它们确实是分解的，而且其中与各个区域相联系的游戏的值可以被计算出来．作为例子，下面图 14.10 和图 14.11 中的围棋局面分别具有像我们的老朋友 1/2 和 * 那样的值． [108]

图 14.12 显示了图 14.1 中排局问题的分解．凝神观赏吧！

在这个例子中，白方先行，赢一目，但是他必须每一步都走对．有一个软件包，它能在这个游戏以及其他归属于其问题集的游戏中以最优策略走子．这个软件并不对棋局发展作预估；它的计算仅由值的基本操作所组成，但它能滴水不漏地走子，因为它有一张准确地记录着棋盘上所有局部区域之值的表．

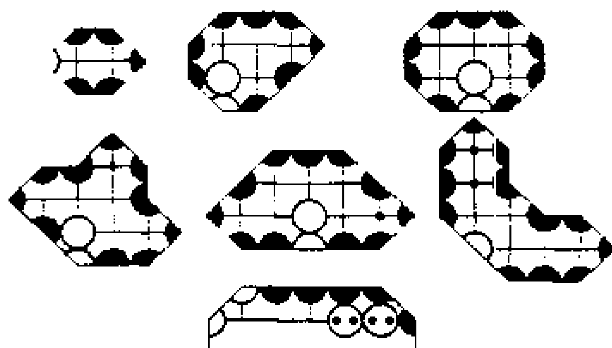


图 14.10

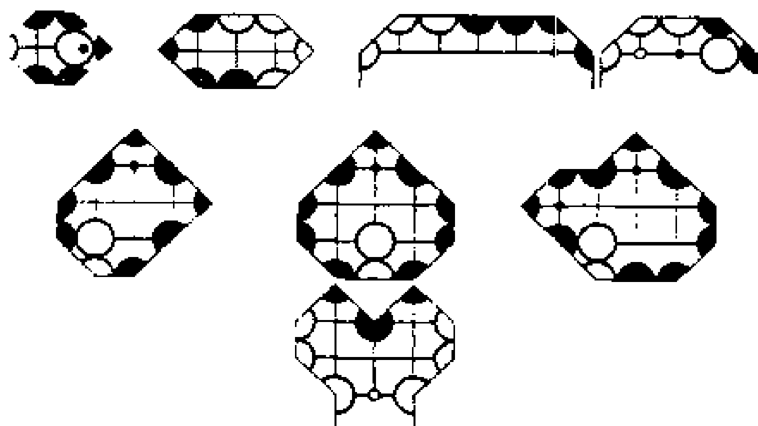


图 14.11 [109]

图 14.12

展　望

接下来一个自然的问题是:我们从这里走向何方?让高手们抓耳挠腮固然是很有乐趣的,但是既然人们知道能做到这一点,那么目的也就达到了.然而还有其他的理由要把这种研究进行下去.伯利坎普指出,职业围棋手已经构想出各种各样十分成熟但多少有点不精确的原则,供他们在下终盘时使用.他相信,利用CGT,这些"经验法则"中有一些可能会得到精确的数学表述.人们可以把这个雄图大略看作是对现代对策论早期工作的一个非常精致的模拟.这里所指的现代对策论早期工作,具体地说就是博雷尔和冯·诺伊曼(von Neumann)在给出关于二

人扑克游戏中讹诈策略[1]的完整数学理论时得到的结果.

一些反思

或许我应该为在前面的段落中公然进行教学表示道歉.最初,我对组合对策论是一无所知,而且,应该很显然的是,我至今也只不过涉及它的表面.然而,这一经历给我留下了对数学本身那种不可名状的多样性的一种压倒一切的敬畏感.《取胜之道》的作者们创造了一个数学仙境.据我所知,在这个学科中还没有其他地方像它那样遥远.想到这一点,我忍不住多少用一些思辨来作为结束.

[110]

常听到有人说数学的所有一切实际上是一个整体,而且随着这个学科的发展,我们最终将清楚地看到每一件事物是如何与其他每一件事物相一致的.在我看来,布尔巴基的整个事业就是基于某种像这样没有明说的假设.那些作者看来一直在努力帮助我们找到"看待事物的正确方法".但是我很想知道这种对终结统一性的信念是否不致于仅为一种一厢情愿的想法.当然,如果从少数几个指导性原则出发,我们能得到对每一件事物的一个理解,那将是十分美好的.不幸的是,现有的证据中没有一件能提示这种情况终将发生.比方说,在超限基数理论一方与偏微分方程数值解这另一方之间,会有多少统一性呢?

在数学之外的科学中,人们也听到关于统一性与多样性的争论.物理学家谈论着他们正在寻找的大统一理论.或许已设计建造的"超级对撞机"[2]将会提供这方面的一些线索,但如果

---

① 原文为bluffing,意思是手中拿的明明是坏牌却仍下大赌注以把对方吓得放弃的策略.《自然杂志》1986年第12期上署名木其的文章《对策论、外交谈判及其他》结合著名的慕尼黑谈判对此作了精彩的介绍.——译注

② 这是指美国于1989年动工建造的"超导超级对撞机".但这项工程从一开始就受到质疑.1993年10月,美国参众两院联席会议以多数票决定中止这一工程.详情可参见《科学》1993年第6期上郁忠强的文章《超导超级对撞机》.——译注

经验有什么指导作用的话，那么这一大堆硬件结果所产生的难题可能比它所解决的还要多．在这种情况下，我猜想我们会继续下去，建造“顶级对撞机”．

我本人的直觉是，数学（或许还有物理学）不会走向统一．它只是不具有兽性．甚至布尔巴基式的努力最终也可能是弊大于利，因为它们把人们的思想强迫纳入一个规定的框架．我预期组合对策论只是自由自在的数学想像所能创造的一个例子，肯定还将有许多其他的例子．我猜想数学非但不会变得统一，相反，它将以完全不可预测的方式继续多样化．这看起来可能是一种在某种程度上有点可怕的想法，但是如果正如我所相信的那样，这就是我们生活在其中的宇宙的本性，那么我们还是应该勇敢地予以面对．

到处是定理

莎士比亚写道，可发现“树木间的谈话，溪流中的文章，石头中的启示，以及每件事物中的益处”[1]．如果他那时喜爱数学，那么他可能还会在某些地方发现定理，或许就在云端．是的，我想具有“数学意识”的人经常会在一些没有预料到的地方一头碰上定理．

举一个例子．我最近想起了一种智力游戏，它过去经常——或许现在仍然——出现在儿童杂志中．它要求你用尽可能少的步数——比方说——从 SHIP（船）走到 DOCK（码头），走法由下面这个可能的解答予以例示．

---

① 此语出自莎士比亚喜剧《皆大欢喜》（*As You Like It*）第二幕第一场中老公爵（Duke Senior）的台词．原文为“…Finds tongues in trees, books in the running brooks, / Sermons in stones, and good in everything.”我国著名莎士比亚作品翻译家朱生豪的译文是：“……却可以听树木的谈话，溪中的流水便是大好的文章，一石之微，也暗寓着教训，每一件事物中间，都可以找到些益处来．”（见《莎士比亚全集（三）》，人民文学出版社，1978年版．）这里因上下文关系，采用了直译．——译注

SHIP, SHOP, CHOP, COOP, COOK, COCK, DOCK.

(这个游戏会导致出现四字母下流词[①]. 这一事实让人想到, 对于那些不把时间花在做更有意义的事上的人来说, 可能还会有更为不堪入目的变化形式.)

于是这里有

船-码头定理 *在这个问题的任何解答中, 一定存在一个词, 它的字母中至少有两个是元音字母.*

当然, 我不是要把它作为一个挑战性问题提给这本杂志[②]的读者们, 而是它可以在其他方面派用处. 正如已经提到过的, 这是出自"日常生活"的定理的一个例子. 而且, 这个结果可以合于理解地被认为是应用的数学, 因为它可用于解决船-码头问题: 它提示用双元音字母词从中间开始而不是从任何一头开始. 还有, 它提供了又一个示例, 这个示例(与哥尼斯堡七桥问题[③]、帕普斯定理[④]等其他示例一起)使人们相信, 数学并不是他们中许多人认为的那样是对数的研究. [111]

然而, 对我来说最为感兴趣的是, 这个结果有可能应用于教学. 显然, 理解这个问题不需要什么数学背景. 在有限的几次实验中, 我曾经试着向一些人陈述这个结果并要求解释它为什么一定是这样(我尽量避免使用"证明"这个词, 看来这个词会立即导致某些人的惊慌). 反应是各种各样的. 从(a)近乎对抗,

---

① 英语中有一些粗俗下流的词是用四个字母拼成的, 一般称为"四字母下流词". 如这个示例中的 cock, 本义是"公鸡", 但又是对男性生殖器的一种粗俗叫法.——译注

② 指刊印本文的杂志《数学信使》.——译注

③ 关于哥尼斯堡七桥问题以及欧拉的解答, 可参见本译丛中《数学娱乐问题》的第4章.——译注

④ 关于帕普斯定理, 可参见本译丛中《近代欧氏几何》的第13章(该书中提到两个帕普斯定理, 另一个在第6章. 但一般说到帕普斯定理, 总是指第13章中那个著名的射影几何定理).——译注

"听着,我离开学校就停止接受考试了",到(b)窘迫,"我们换个话题吧",到(c)微光闪现,"这很显然,因为 Ship 和 Dock 中的元音字母处在不同的位置"(你的路子走对了),到(d)"哎呀,我完全明白这一点,但我不能解释"(这种说法我以前在什么地方听到过?),到(e)"这是因为每个词都必须至少含一个元音字母"(很好,你几乎要成功了!),到(f)完成证明,以及表现更好的(g)因看到这些逻辑片段是怎样结合起来的而感到快乐.当然,正是(g)让我们不断前进并使数学的实践成为一个如此有回报的职业.作为教师,我们的部分任务就是调动我们的学生离开情况(a)而走向情况(g).不幸的是,据我所知,至今还没有人想出怎样做到这一点.或许我们所能做的最好的事就是在云端搜寻更多的定理.[112]

第15章 再说悖论——知识游戏

我提出下面的例子，以再次引入悖论这个话题（见第2章）.

有人把数 a 和 b 分配给你我二人，它们是相邻的正整数，比方说 $b = a + 1$. 我们每人都知道自己的数，但不知道对方的数. 一台报时器每隔10秒钟发出一次嘟嘟声响，这时如果我们中有哪个知道了对方的数，他就必须在这嘟嘟声响结束后立即宣布这一点，于是游戏结束.

这个游戏看起来颇为荒唐. 这间隔发生的嘟嘟声响会对局中人有什么帮助呢？例如，假设我的数是10，你的数是11. 那么我们都知道在第一次声响后我们中没人会作宣布，因此当这样子继续进行下去的时候，我们的情况似乎不会比刚开始时有所好转. 然而，我们有

定理 *在绝对正确的玩法下，掌握 n 的那个人将在第 n 次声响后宣布对方的数是 $n+1$.*

对 n 用数学归纳法. 如果我的数是1，我便知道你的数是2，于是我将在第一次声响后予以宣布. 现在假设我的数是 n. 如果你的数是 $n-1$，则根据归纳法假设，你将在第 $n-1$ 次声响后作出宣布，因此一旦你没有这样做，我就知道你的数是 $n+1$. [113]

虽然这个证明是正确的，但是仍令人迷惑不解. 一个我们都知道会发生的没人作宣布的情况，怎么会导致游戏对局形势的变化呢？为了弄明白这一点，假设我的数是3，你的数是4. 那

么我们事先都知道在第一次声响后不会有人作宣布．然而，我不知道你知道这一点，因为就我所能知道的，你的数也可能是2，假如是这种情况，你会想到我的数可能是1，从而推测我可能马上要作一个宣布．因此，这第一次声响给了我新的信息①．如果我们的数是4和5，那么我确实知道你知道在第一次声响后不会有人作宣布，但是你却不知道我知道这一点，如此等等．

这是已开始被称为**公共知识**(common knowledge)游戏或迷题的智力游戏的一个例子．这里还有一个．有一个学校，全是聪明伶俐的儿童．一天早上，在一个有20个孩子的班级里，有14个孩子竟脏着脸来上学．老师说，“我们来玩一个游戏．你们大家可以相互看看各人的脸．这台报时器将每隔10秒钟响一次．如果你知道了你自己的脸是脏的，就请在声响后举手．”游戏开始了，但过了几分钟，什么也没发生．于是这位老师说，“我看我还得给你们一个提示．你们当中至少有一个人的脸是脏的．”因为每一个孩子都至少能看到13张脏脸，这似乎并不是什么重大的新闻．然而，当这游戏重新开始，到第14次声响后，所有的脏脸孩子都举起了手．这个证明留给读者，仍然是用数学归纳法，这次是对脏脸的个数作归纳．

从那个提示中人们得到了什么新的信息？让我们考虑只有你和我是脏脸的情况．在提示之前，我知道你不会在声响后作宣布．但在提示之后，就有了你会作宣布的可能性；具体地说，是在我的脸是干净的假设下．如果还有第三张脏脸，那么我们都知道在第一次声响后没人会作宣布，但是我不知道你们知道这一点——如此等等．

关于这个主题的最数学的变体是康韦和佩特森发明的一个游戏[1]．一个房间里有N个人，每个人的额头上都写着一个非

---

① 具体地说，这个信息是：由于没人作宣布，因此如果你的数是2，那么你就排除了我的数是1的可能性，从而你将在第二次声响后宣布我的数是3．—— 译注

负数，第 k 个局中人的数是 a_k. 此外，还有不多于 N 个的不同正数 A_k，写在一块黑板上，其中有一个是那些 a_k 之和. 这些数不一定是整数. 每一位局中人都看到了除自己额头上的数以外的所有数. 我们还是使用那台每隔 10 秒钟响一次的报时器. 这里的目的是，看看要响多少次才会有某位局中人知道他自己的数，或者说，知道哪一个 A_k 是真正的和，这是一回事. 康韦 - 佩特森定理断言，这个游戏总会结束.

我们又有了一个悖论. 例如，假设在这样的一个三人游戏中，所有小写的 a 是 2，而大写的 A 是 6，7 和 8. 于是每个人都知道，自己的数至多是 4，而对手们所看到的两个数，其和至多为 6. 这样，那 3 个 A 中的任一个都可能是真正的和，因此在第一次声响后不会有什么反应.

关于康韦 - 佩特森定理的最令人意外的事情或许是，它的证明极其简单，条件是你要用正确的方式来解读. 不过，在介绍这个证明之前，让我们先考察一下局中人是两个，而黑板上的两个数是 $A < B$ 的情况. 在这样的 (A, B) 游戏中，当且仅当一对额头上的数 (a, b) 满足 $a + b = A$ 或 $a + b = B$ 时，我们就说 (a, b) 是可能的. 于是由可能的数对 (a, b) 所组成的集合就是平面上的一对斜线段. 见图 15.1.

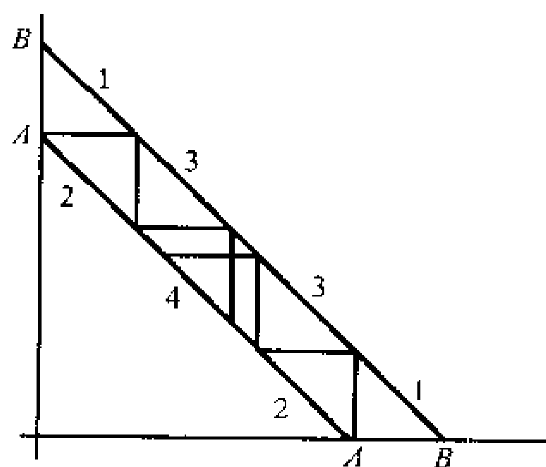


图 15.1

这样，在一次声响后就结束的游戏正是那些其中有一个元素 x 大于 A 的数对，这是因为看到这样一个 x 的局中人将知道真正的和是 B. 这些数对就是图 15.1 中线段 B 上标着 1 的两条线段. 排除掉这些包含一个大于 A 的数的数对后，如果有一个数对包含一个小于 $B-A$ 的数，那么相应的游戏将在第二次声响后结束，这是因为看到这个数的局中人将知道真正的和是 A. 这些数对就是线段 A 上标着 2 的线段. 在这种方式下，每次声响都让人们可以截去那两条斜线段中某一条两端的线段. 从图 15.1 人们看到线段 A 在 4 次声响后全被排除；因此，比方说假设 $a=b=B/2$，那么局中人在第四次声响后将会知道这一点.

[114]

在给出一般的证明之前，我们要引进另一个令人意外的事实，它是由拉斯里(J.M.Lasry)、莫雷尔(J.M.Morel)和索利米尼(S.Solimini)[2] 发现的. 甚至当黑板上的数比局中人的个数还要多时，康韦－佩特森游戏也可能结束. 例如，我们将证明，当局中人是两个，而黑板上的数是 5，8 和 15 时，这个游戏至多在 10 次声响后结束. 表 15.1 给出了这个游戏的可能的数对①.

5	8	15
(0,5) 5	(0,8) 6	(0,15) 1
(1,4) 9	(1,7) 8	(1,14) 1
(2,3) 3	(2,6) 2	(2,13) 1
	(3,5) 4	(3,12) 1
	(4,4) 10	(4,11) 1
		(5,10) 1
		(6,9) 1
		(7,8) 7

[115]

表 15.1　可能的数对

① 从这里开始，作者似乎在“偷梁换柱”. 原先说得很明白，这些数不一定是整数，但下面的分析都假定这些数是整数. 其实，有理数的情况显然可用“通约”的方法化归为整数的情况；而无理数的情况则可用有理数的情况来“逼近”. ——译注

同往常那样，在着手分析时，首先找出使得游戏在一次声响后即结束的数对，其次找出那些使得游戏在两次声响后结束的数对，如此等等．表 15.1 中放在每个数对括号后面的数字，表示相应的游戏到结束时所响的次数．首先，如果有一个数大于 8，那么看到这个数的局中人将知道真正的和是 15，从而就会在第一次声响后予以宣布．注意从 9 到 15 这些数，它们正是那些仅在一个可能的数对中出现的数．删掉这些数对后，我们看到只有一个数对(2,6) 包含着一个 6，因此如果有一位局中人看到一个 6，他将知道真正的和是 8，并在第二次声响后作宣布．排除掉这个数对，我们发现(2,3) 是仅存的包含 2 的数对，因此接下来就把它排除掉，如此等等．这规律再简单也不过了．在每一步，找出所有只被一个数对所包含的数并排除掉这些数对．要响 10 次的游戏就是每位局中人的数都是 4 的那个．

另一方面，如果上面例子中的 8 改为 9，我们将证明这个游戏不可能结束．同前面一样，我们看到，当且仅当有一个数大于或等于 10 时，这个游戏将在第一次声响后结束．然而，到这里我们却无法前进了，因为如表 15.2 所示，每个数都在余下数对的两个数对中出现．因此无论局中人看到什么数，他们都将无法就真正的和得出任何结论．

5	9	15
(0,5)	(0,9)	(0,15) 1
(1,4)	(1,8)	(1,14) 1
(2,3)	(2,7)	(2,13) 1
	(3,6)	(3,12) 1
	(4,5)	(4,11) 1
		(5,10) 1
		(6,9)
		(7,8)

表 15.2　可能的数对

在[2]中,作者们对两个局中人、三个黑板数的游戏进行了完全的分析. 假设 $A_1 < A_2 < A_3$,那么这样的游戏会结束的充分必要条件是,存在整数 p 和 q,使得 $A_1 < p(A_2 - A_1) + q(A_3 - A_2) < A_3 - A_2$.

明白了上面第二个例子,我们现在可以给出关于 N 个局中人、N 个黑板数游戏的康韦 - 佩特森定理的那简单得出人意料的证明了. 关键的想法是用反证法证明. 不是去证明游戏必定会结束,而是证明对这种游戏来说不结束是不可能的. 就像在上面那个例子中那样,对于一个不能结束的游戏,一定会在某次声响后遇到不可能再排除任何可能的 N 元组的情况. 这种情况怎样才会发生呢? 答案由一个关于向量空间的简单定理给出.

[116]

对于一个 N 维向量的集合 S,如果对 S 中的任何一个元素 $\boldsymbol{a}$ 以及任何一个指标 i,存在着 S 中的一个向量 $\boldsymbol{a}'$,它与 $\boldsymbol{a}$ 只是第 i 个坐标不相同,则我们称 S 为**模棱两可的**(ambiguous). 如果一个游戏进行到某一步,余下的 N 元组所组成的集合是模棱两可的,那么每一个局中人都不能从下一次声响中得知什么信息:对于他自己的数,至少总有两个可能的值.

给出一个向量 $\boldsymbol{a}$,让我们记 a_0 为 $\boldsymbol{a}$ 的坐标之和.

引理 如果 S 是一个 N 维向量的模棱两可的有限集,那么坐标和 a_0 的集合一定至少包含 $N + 1$ 个元素.

对于 $N = 1$,结论唾手可得. 现在选择这样一个元素 $\boldsymbol{a} = (a_1, \cdots, a_N)$,其中的 a_1 对 S 中所有元素来说是最小的. 并令 S' 是所有这样的 $N - 1$ 维向量 $\boldsymbol{x} = (x_2, \cdots, x_N)$ 的集合,它们使得 $(a_1, x_2, \cdots, x_N)$ 在 S 中. 那么 S' 也是一个模棱两可的向量集合. 因此根据归纳法假设,S' 中的向量至少有 $(N - 1) + 1 = N$ 个不同的坐标和. 这样,形如 $(a_1, x_2, \cdots, x_N)$ 的向量至少有 N 个坐标和. 令 $(b_2, \cdots, b_N)$ 是 S' 中的坐标和为最大的一个向量. 根据 a_1 的选择方式以及模棱两可性的定义,存在 $a_1' > a_1$,使得向

量$(a_1', b_2, \cdots, b_N)$在 S 中,但是这产生了一个坐标和,它大于任何一个已考虑过的坐标和,也就是说,它是第 $N+1$ 个坐标和.

我们指出,$N+1$ 是"尽可能最好的",这可以用坐标仅为 a 或 b 的所有向量的集合作例子来予以说明.在这种情况下,坐标和仅依赖于 b 的个数,它可以在 0 到 N 中变化.

作为一个练习,读者可以试着证明额头上的数分别为 2,2,2,而黑板上的数为 6,7,8 的三人游戏将在响了 15 次后结束.

参考文献

1. J.H.Conway and M.S.Paterson, *A Headache-Causing Problem*, in privately published papers to W.Lenstra on the occasion of the publication of his *Euclidische Getallenlichamen*
2. J.M.Lasry, J.M.Morel, and S.Solimini, On knowledge games, *Revista Mathematica de la Universidad Complutense de Madrid* 2(2/3)(1989).

[117]

[118]

第16章 三角形与计算机

引 言

第6章论及的主题是在计算机帮助下得到的几何学发现，特别是金伯林的关于三角形"中心"的某些工作，这里的所谓中心，就是诸如形心、外心、垂心之类的点. 金伯林定义了91个这样的点，并且通过数值计算上的探究，发现在这些点中间存在着(或者我应该说，看来是存在着)极大量的共线现象. 这些经验性的结果随即可以被证明是正确的，用的仍然是计算机. 但是这次是使用符号计算而不是数值计算：把给出的中心用"三线坐标"表示成边长 a，b 和 c 的函数，有关的计算就变成了这样一件事——证明这些符号的某个相应的行列式为零. 这是一项十分理想地适合于如 *Mathematica* 或 *Maple* 这样的程序软件来完成的工作.

计算机技术的第三种可能的应用，是一种对几何学问题特别自然的应用，它当然就是计算机制图学. 这里介绍的三个例子，其中第一个显示了这种计算机生成的图形是怎样在经典的欧氏平面几何中导致十分令人惊奇的新结果的. 与此相反，我[119]们的第二个例子虽然更为初等，但它处理的问题却完全不同于我所知道的任何几何学论著所论述的任何问题. 在第三个例子中，在实验阶段，我们用铅笔和直尺代替了计算机，但那"画龙点睛的妙着"还是由金伯林的数值实验给出的.

西姆森线之舞

对于一个给定的三角形，令 S 为其外接圆. 从 S 上的任一点 P 向这三角形的三条边作垂线.

定理 这三条垂线的垂足共线.

这条轨迹称为这个三角形关于 P 的**西姆森线**(Simson line)(见图 16.1). [据考特(N.A.Court)在《大学几何》(*College Geometry*, New York: Barnes & Noble, 1925 年, revised 1952 年)中的说法，人们错误地把这条线的发现归功于西姆森(Robert Simson, 1687—1768)，其实，这条线是华莱士(William Wallace)在 1799 年发现的.]

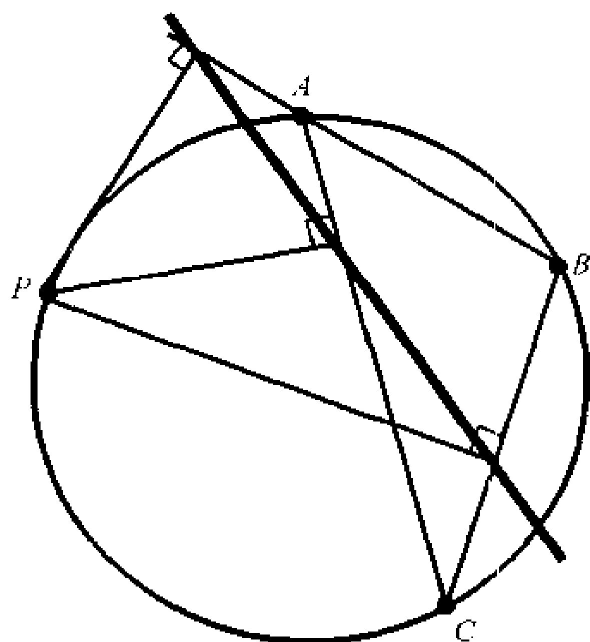


图 16.1

这个三角形的三条(延长了的)高和三条延长边也是西姆森线. 要得到作为西姆森线的高，就将 P 置于 $\triangle ABC$ 的某个顶

点．要得到作为西姆森线的边，比方说要得到 BC，就将 P 置于 A 的对径点．(一画出图，你就会看到这些结论的简单证明．)

接下来，给出一个圆内接四边形 $ABCD$，这就决定了四条西姆森线．它们是：$\triangle ABC$ 关于 D 的西姆森线、$\triangle BCD$ 关于 A 的西姆森线，等等．如图 16.2 所示．

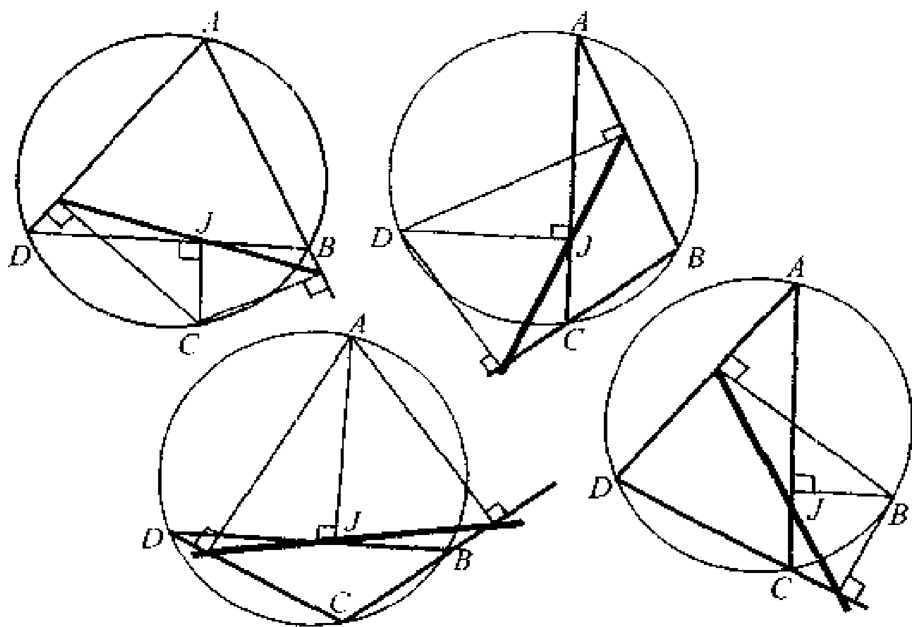


图 16.2

考察了一些像图 16.2 那样的图形后，约翰逊(Dennis Johnson)注意到这四条西姆森线看来是共点的，而且他证明了确实如此．这是一个已知的结果(见考特的《大学几何》，定理304)．不过，我们还是把这个交点称作那已给四边形的 J 点．至此我们还没有用到什么计算机，但是，约翰森接着研究了当 $\triangle ABC$ 保持固定而第四个点 P 围着其外接圆运动时 J 点的轨迹．他用一个计算机程序产生了如图 16.3 那样的图形．

当 P 围着那原先给出的圆运动时，J 点看来是围着这个三角形的九点圆运动．这个九点圆很容易识别：它就是经过这三角形三条边中点的那个圆(另外六点是高的垂足和连接顶点与

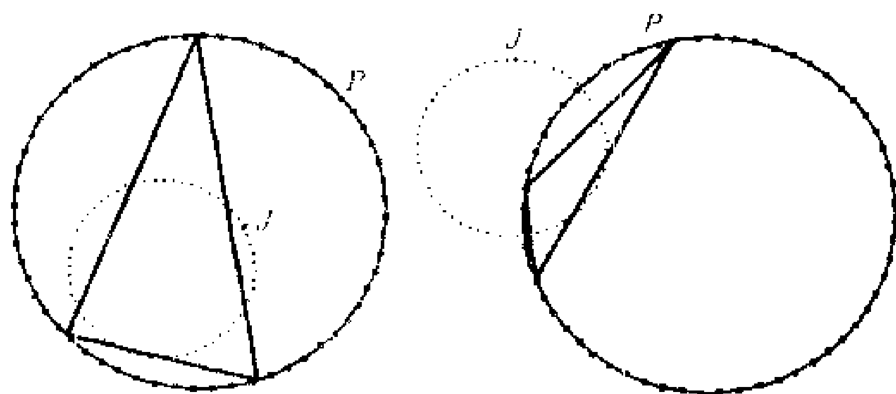


图 16.3

[121]

垂心的线段的中点)．而且，J 点围着这九点圆运动的方向和角速度都与 P 相同．约翰逊找到了所有这些性质的证明．不幸的是，他对这些结果的评论在1991年的奥克兰大火①中遗失了．不过，如果你知道关于九点圆的一些事实，这证明还是不难的．

下一个问题：对于一个给定的三角形，当 P 围着它的外接圆运动时，全体西姆森线组成的族是个什么样子？

约翰逊编了一个程序，它可对任意选取的三角形生成"西姆森族"的图象．值得注意的是，所有这些线族的样子都相同，它们与三角形的形状无关。更准确地说，所有的线族看来是全等的，只是在它们相对于那给定三角形的位置和方向上有所不同．这些族总有一个包络，看上去好像是一个"三尖点内摆线"，这是当一个圆沿着半径为其三倍的圆内周滚动时它圆周上一点的轨迹．图16.4给出了两个例子，一个是关于锐角三角形的，一个是关于钝角三角形的．接下来是什么呢？

① 1991年10月20日，美国加利福尼州奥克兰和伯克利附近的风景山区发生大火．25人丧生，150人受伤，烧毁房屋约3000幢，财产损失达15亿美元．——译注

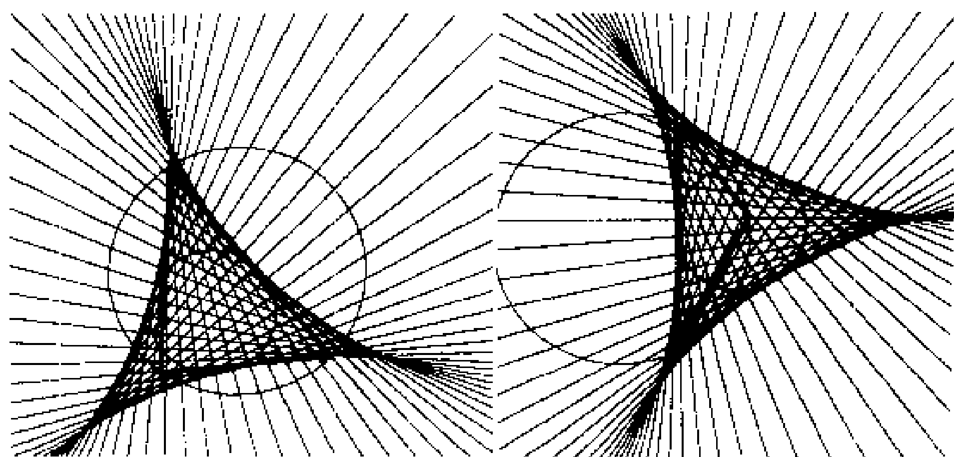


图 16.4

回答还是与九点圆有关．当那个图形程序经修改后把这九点圆也画出时，如图 16.4 所示，我们清楚地看到，它正是西姆森族的包络的内切圆．根据这个观察结果，约翰逊完成了这一舞蹈的全部编排设计．芭蕾舞女演员西姆森小姐可以被想像成拿着一根长长的平衡杆，就像走钢丝的杂技演员所用的那种．它代表着西姆森线．当 P 围着这个圆运动时，在每个瞬间，这位芭蕾舞女演员位于相应的 J 点，并由此用某种相同的角速度 ω 围着这九点圆——比方说以顺时钟方向——运动．同时，她必须以逆时钟方向用一种慢动作的方式作自身旋转，角速度是 $\omega/2$．于是那平衡杆就扫遍了西姆森族．这给出了全部的故事．包络上的那三个尖点出现在平衡杆通过九点圆圆心的时候．我们把这样的一条线称作**尖点线**．假设在开始时平衡杆通过这圆心，并令这条尖点线为参考线．那么当从圆心到 J 点的半径线转过一个角度 Θ 时，平衡杆将与这半径线形成一个 $(3/2)\Theta$ 的角度．因此 J 点转一圈，平衡杆就通过那圆心三次，从而形成三条尖点线．而那包络最后被证明是这九点圆（半径为 1/2）在一个半径为 3/2 的圆内滚动而得到的内摆线．

这故事还未完全结束．请注意这西姆森内摆线是初始三角

形的一个不变量. 因此会产生一些自然的问题,例如,对一个给定的三角形,怎样求出它的——比方说——尖点线. 最后又是一个意外:一般来说,这些线并不是尺规可作的. 约翰逊巧妙地证明了,如果尖点线是可以作出的,那么人们就可以三等分任意角,而我们知道,这件事用圆规和直尺是不可能做到的.

有理角的构形

作为背景知识,读者应当去阅读本书附录 3 中那篇由马查多(Armando Machado)撰写的引人入胜的文章《初等几何中的十九个问题》. 在图 16.5 中,我们重现了马查多问题中的第一个:给出一个等腰三角形,顶角为 20°、直线 a 和 b 分别与底边成 60° 和 50° 角,请确定未知角 γ. 建议读者花一点点时间试着解答一下这个问题,这样便可体会到它的难度. 写出 γ 所必须满足的各种三角方程并不难. 马查多用数值计算的方法(只用一个袖珍计算器!)解出其中一个方程,结果很意外地发现,在计算器的精确度范围内,γ 为 80°. 这个发现激发了他的兴趣,于是他用软件 *Mathematica* 进行了数值搜索,又找到了几十个这种"有理构形"的例子,即其中所有的角度都是 π 的有理数倍. 对这些例子按解决方法进行归并,马查多总结出了 19 种不同的情况. 约翰逊用计算机对一种不同的类型进行了搜索. 简短地说,在这种类型中,直线 a 和 b 与三角形的边相交时可以不一定交于线段内部. 这就又找到了 250 多个有理构形的例子. 这些例子中,有一些属于各种无穷的线性族,而另一些则看来是孤立的. [123]

当然,数值计算上的任何发现,不管它多么令人相信,都不能构成一个这些构形确实存在的证明. 因此,马查多的 19 个问题就是要求找出这些构形存在的证明. 就像通常在几何学中那样,有两种证明方法,即综合法和代数法.

图 16.5 中的右图给出了马查多最初那个例子的一种综合法证明. 画直线 AB 和 BC,其中 AB 与底边成 20° 角. 然后证明

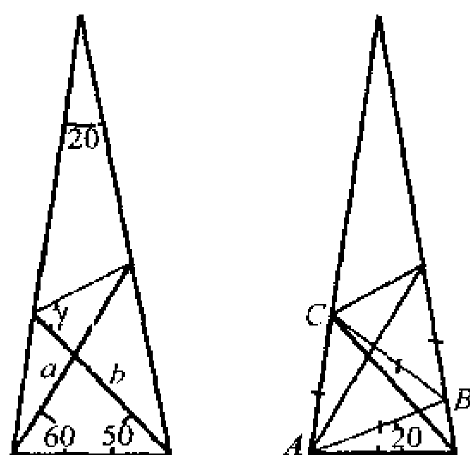


图 16.5

所有标上记号的线段都相等．据此并据已知的角度，就可得到 γ 的角度．事实上，拉马尔霍(Margarita Ramalho)成功地找到了顶角为 $20°$ 的六个马查多构形的综合法证明．然而，令人惊奇的是，每一种情况都需要一种不同的证明和一套不同的多达四条的辅助线．因此，要想对所有这些已发现的例子——有些例子中的角度 $(k/n)\pi$ 还有着比方说能被 7 整除的分母——都找到初等的综合法证明，看来是一件没有希望的工作．

人们想要的是一种不变的程序，一种算法，用它可判定一个具有给定角度的有理构形是否存在．既然未知角 γ 被图 16.5 中的给定角所唯一确定，显然这四个角必须满足某个三角方程．约翰逊把图 16.6 中所示的角度选作参数．这些角度之间的关系可用各种各样的方程表示．例如，约翰逊找到了这样一个方程：

$$(\sin^2 T)\sin(B+E-A)=\sin A\sin B\sin E.$$

把三角函数转换成复指数函数是一种有效的方法．令 $\alpha=e^{2iA}$、$\beta=e^{2iB}$，$\tau=e^{2iT}$、$\varepsilon=e^{2iE}$，我们得到

$$\begin{aligned}&(\tau-1)(\tau^{-1}-1)[(\alpha^{-1}\beta\varepsilon)-1]\\&=(\beta-1)(\varepsilon-1)(\alpha^{-1}-1).\end{aligned}\tag{1}$$

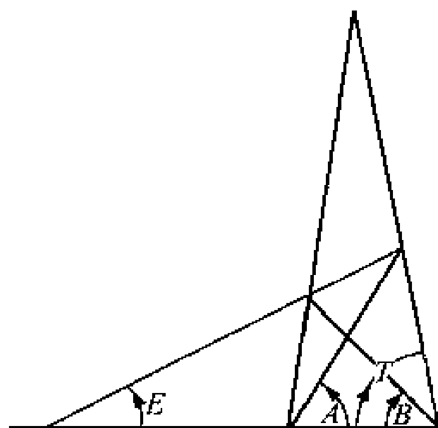


图 16.6

[124]

现在,如果A,B,T,E这几个数是π的有理数倍数,那么$\alpha,\beta,\tau,\epsilon$就是单位根[①]. 例如,在最初那个问题中,$A=(1/3)\pi$、$B=(5/18)\pi$,$T=(4/9)\pi$,$E=(1/6)\pi$. 它们有最小公分母18. 取$\xi$为18次本原单位根,我们得到$\alpha=\xi^6$,$\beta=\xi^5$,$\tau=\xi^8$,$\epsilon=\xi^3$. 把它们代进方程(1),就得到了多项式方程

$$\xi^{17}+\xi^{15}-2\xi^{12}+2\xi^8-\xi^5-\xi^3+\xi^2-1=0. \qquad (2)$$

为证明ξ满足这个方程,我们验证它能被18次本原单位根的分圆多项式$\xi^6-\xi^3+1$所整除. 事实上,方程(2)可因式分解成

$$(\xi^6-\xi^3+1)(\xi^{11}+\xi^9+\xi^8-\xi^6-2\xi^3+\xi^2-1).$$

这种方法显然具有一般性. 约翰逊用这种方法发现了他的所有那些有理构形. 对每一个正整数N,令A,B,T,E取所有小于N的正整数[②]. 然后把ξ的相应幂代入方程(1),并进行检验,

---

① 即方程$x^n-1=0$的复数根,它们可表示成$\cos(k/n)2\pi+i\sin(k/n)2\pi=e^{2i(k/n)\pi}$,$k=0,1,\cdots,n-1$. 其中有一种单位根$\xi$,它使得$\xi^0,\xi^1,\cdots,\xi^{n-1}$就是$x^n-1=0$的所有复数根. 此即$n$次本原单位根. 设$\xi_1,\xi_2,\cdots,\xi_m$是所有的$n$次本原单位根,则可证$(x-\xi_1)(x-\xi_2)\cdots(x-\xi_m)$是一个整系数不可约多项式,此即分圆多项式. 详请可参见《代数数论导引》(张贤科著,湖南教育出版社,1999年版).——译注

② 此处表述似有误. 应为"令a,b,t,e取所有小于N的正整数,并令$A=(a/N)\pi$,$B=(b/N)\pi$,$T=(t/N)\pi$,$E=(e/N)\pi$".——译注

看它是否能被 N 次分圆多项式所整除. 这种方法不同于数值搜索,它也是证明构形的存在. 使用一种类似的技巧,鲁滨逊同样验证了马查多文章中所有构形的存在.

尽管现在我们有了大量的数据,但有理构形的集合的构造仍然是一个难解的谜. 约翰逊对 N 为从 8 到 36 的所有偶数都找到了有理构形,但当 N 为这个范围内的奇数时却一个也没有(除了像 $\alpha = \beta$ 那样的平凡情况). 或许当 N 为奇数时并没有非平凡的有理构形,如果真是这样,为什么? 还有一个谜:马查多那个最初的例子源于何处?

一些后来的消息

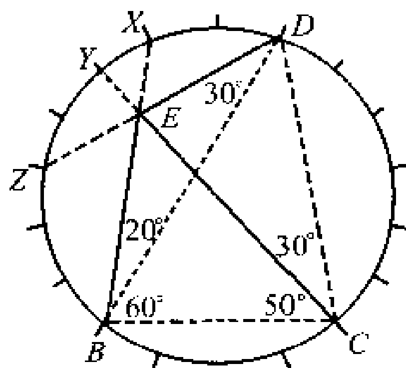
我非常感激唐·查克里安(Don Chakerian),他向我提供了与上述两项内容密切相关的信息. 关于"西姆森线之舞"这一节,查克里安写道:

> 戴维·凯(David Kay)在他的《大学几何》(*College Geometry*, Holt, 1969 年)中提到(第 248 页),那些西姆森线生成的内摆线显然是由斯坦纳(Jacob Steiner)首先发现的. 凯给出了继洛克伍德(E.H.Lockwood)的《西姆森线及其包络》(Simson's Line and it's Envelop, 载《数学公报》(*Math. Gaz.*) 37(1953), p.124—p.125)之后的一个进展情况.

第二个信息是关于"有理角的构形"这一节的. 我在上面这一节结尾处求问有关马查多文章中第一个问题之起源的信息,关于这个问题,他说道:"……一定相当有名,因为它在数学圈里一再被提到." 多亏查克里安和瓦贡(Stan Wagon),我能够循迹找到了看来是这个例子起源的地方,那是《数学公报》1922 年第 11 卷上由兰利(E.M.Langley)提出的一个问题. 更为有趣的是这样一件事:其实存在着由"兰利问题"引出的一小批文献.

事实上，马查多文章和上面这一节中的所有结果，甚至还有更多的结果，原来都已为人们所知．我不打算给出完整的文献目录 [125]（最好查的参考文献看来是在《数学公报》1978年第62卷的第174至183页），但是这最后的故事颇为吸引人．

有一个问题可称为广义兰利问题：找出所有的"有理四角形"，即所有的其六条边中任何两条的夹角都是 π 的有理数倍的完全四角形[①]，并对它们进行分类．这个问题于1978年被蒙斯基（Paul Monsky）完全解决．他证明，除了明显的例子（例如，一个具有有理角的三角形加上它的角平分线[②]）外，这些解由120个单参数族和1830个孤立情况组成．蒙斯基的手稿大约有30多页，但从未发表，因为他的结果原来早在40年前（！）就被荷兰几何学家博尔（Gerrit Bol）发表在一篇论文（荷兰文）《有奖征答第17题的答案》（Beantwoording van Prijsvraag no 17，载《数学新档案》（*Nieuw Arch. Wisk.*）（2）18（1936），p.14—p.68）上了．



把这个博尔－蒙斯基结果与上面所述的一些情况结合起来，不难证明这个四角形问题相当于问一个正 n 边形的三条或更多条对角线在什么时候共点．在上图中就最初的兰利问题对此作了例示．最终证明，这种共点性在 n 为奇数的情况下不会

① 完全四角形是由平面上一般位置的四个点以及连接它们的六条直线所组成的图形，这与四边形的概念不同．——译注

② 即这个三角形的顶点和内心及它们间的连线组成的完全四角形．——译注

出现,而且除了显然的情况外,它只是当 n 能被 6 整除时才出现. 这就肯定了上面讲过的约翰逊在他的计算机搜索中所观察到的结果. 其他还有一些令人感兴趣的结果,比方说博尔找到了可使 4 条、5 条、6 条和 7 条对角线共点的 n 值,并且证明了只有这几种可能.

三角形中的三角形

好吧,孩子们,今天我们要做一个几何实验. 现在,我明白你们都知道怎样编写画三角形和垂线的计算机程序,但今天我要给你们演示一种不同的方法. 你们只需要一支铅笔(你们还记得它们,是吗?)和一把能画直角的直尺. 你们每人都有一张纸,上面都画着一个同样的三角形,边为 a,b,c. 现在请听我的指令. 我要你们在自己的纸上任意选择一个点,并从此点向直线 a 作垂线. 作好了吗? 从垂足出发,向直线 b 作垂线,从这个新的垂足出发,向直线 c 作垂线,然后作垂线返回直线 a,如此不断地作垂线,一轮又一轮地作下去. 你们注意到了什么? 对了,卡尔 · 弗里德里克,没过多久,你们就在你们原先那个三角形内部的同一个三角形上兜圈子了. 而且,它看起来是把原先那个三角形的每条边转过来并缩小一下而得到的(见图 16.7).

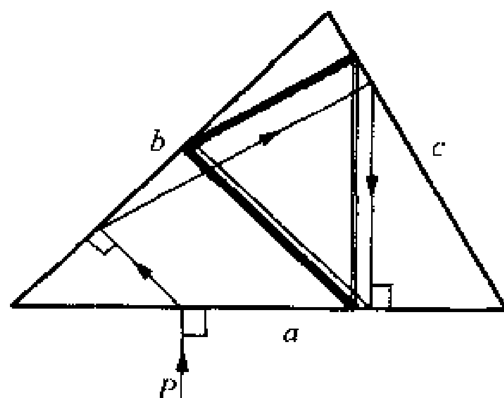


图 16.7

请注意,尽管你们各人选择了不同的出发点,但你们得到了同样的三角形!这是怎么回事,亨利?一种"压缩映射"?嗯 …… [126]

好吧,现在我们要试做一些稍稍不同的事.从一个新的出发点开始,但这次你们不是以从 a 到 b 再到 c 的顺序作垂线,沿另一条路,首先是向 a 作垂线,然后向 c,再后向 b,因此你们将以相反的方向兜圈子.你完全正确,索尼亚,这次我们得到了一个不同的三角形,但它实际上就是我们前面得到的那个,只是它做了个倒立(见图16.8).

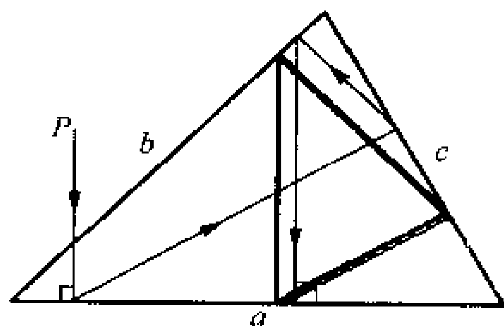


图 16.8

这些现象是胜浦英文(Hidefumi Katsuura)发现的.把原先那个三角形作 ±90° 的旋转,这两个极限三角形便与其位似①,这件事是显然的.但胜浦证明了它们实际上是全等的.它们有着同一个外接圆,并且通过以这个圆的圆心成对称的方式联系在一起.还有,原先那个三角形在这个对称变换下的象包含了这两个极限三角形的六个顶点.所有这些都表示在图16.9中.

到这一步计算机还没有登场.不过,请注意图16.9中的对称中心 P 是原先那个三角形的一个不变量.于是产生了一个自然的问题:它是一个什么心?是形心,垂心,内心,外心,还是九

① 原文作 similar(相似),但说相似就不必把三角形作旋转.—— 译注

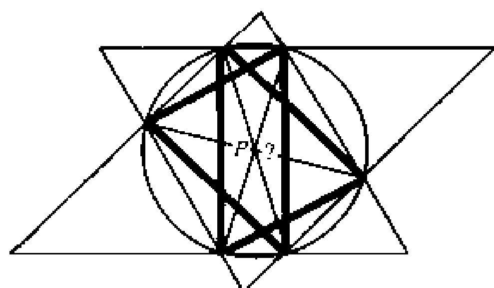


图 16.9

[127]
点圆圆心？为了回答这个问题，我做了一件理所当然的事，即迅速写了一封电子邮件，把有关构造描写了一下，发给埃文斯维尔的金伯林，他就是前面提到的那位研究三角形各种中心的世界级专家．金伯林稍稍做了一些数值实验后，就在发回的电子邮件中告诉我，这个点不是上述那些中心中的任何一个，但是它的坐标与**类似形心**（symmedian point）或称**勒穆瓦纳点**（Lemoine point）的坐标基本相符，其精确度达到14位数字！

什么！你从来没听说过这个类似形心？类似形心就是形心的共轭点．共轭点？是的，选择任何一个不是三角形顶点 A，B，C 的点 P．将直线 PA，PB，PC 分别作关于角 A，角 B，角 C 的角平分线的镜射．这样得到的直线是共点的（定理），它们的交点 P'，就是 P 的共轭点．

知道了——或者我应该说猜想到——P 是勒穆瓦纳点，人们就能解析地证明它确实是如此，就像克利福德·加德纳所做的那样．或者更好的做法是，回到考特的那本《大学几何》，翻到第10章，"近期的三角形几何"（Recent Geometry of the Triangle"、B节，"勒穆瓦纳几何"（Lemoine Geometry）（勒穆瓦纳（Emile Lemoine），1840—1912），定理593．由这个定理很容易得到胜浦的结果．图中的圆称为**勒穆瓦纳的第二个圆**（Lemoine's second

circle),而那三条直径则称为**勒穆瓦纳反平行线**(Lemoine antiparallels). 不过,胜浦对他的结果给出了直接的初等证明,因此不需要用到 19 世纪的数学知识.

补遗:拼图悖论

我们的第一节专门展示了用计算机制图学来解决几何问题的威力. 我们在结束时便展示一下滥用计算机制图学来提出谬误解答的威力.

图 16.A 1 被认为是卡罗尔的杰作,这是一个著名的数学把戏. 它宣称几何图形被分割成有限块以后其面积不一定保持不变.

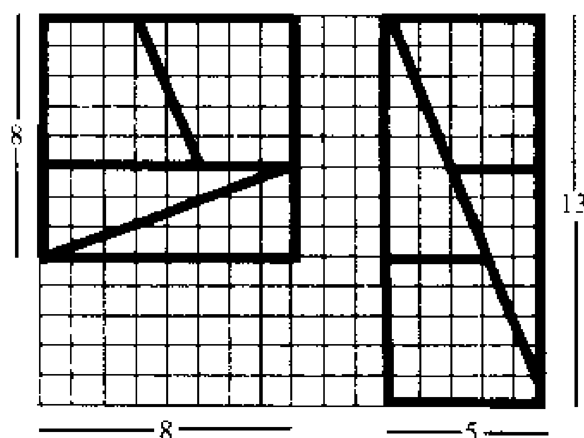


图 16.A 1

128

以前没有见过这个的读者应当试着找出其中的捣蛋鬼. 就是这个一般的原理,已经被用来创作诸如此类的图形分割悖论,例如图 16.A 2 中的那个.

最近,布雷特(Jean Brette) 发明了一大批关于这个主题的精心构思的变化形式. 多年来,布雷特一直是"发现宫"(Palais de la Découverte)—— 即巴黎科学馆 —— 数学方面的负责人.

由图 16.A 3 和图 16.A 4 所表明的例子,实际上是合六个悖

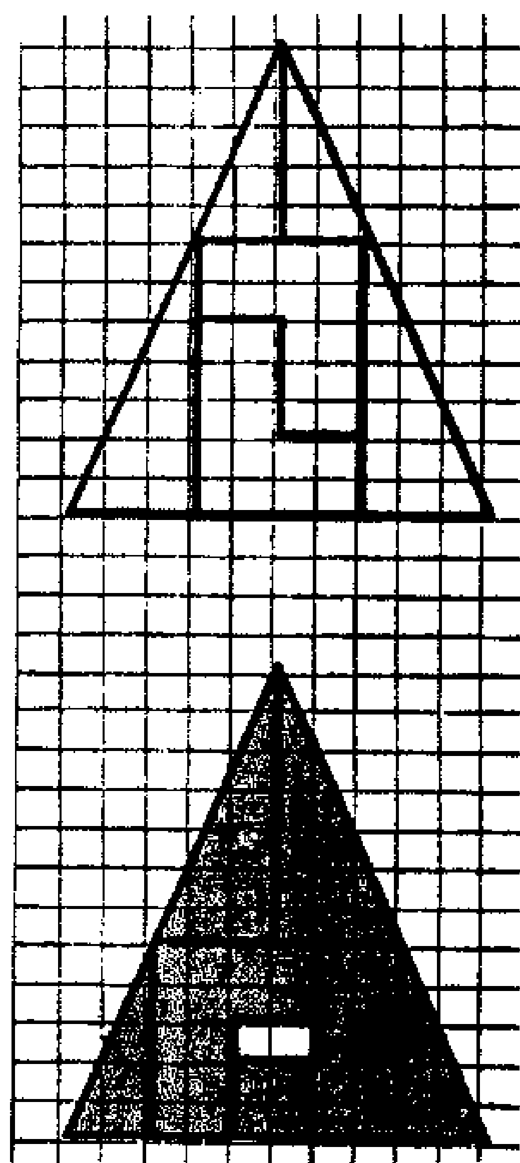


图 16.A 2

129

论而为一个的．利用图 16.A 3 所示的 10 个拼块的各种子集，人们可以用六种不同的方式拼成一个 9×16 的三角形，这六种方式对应于这个大三角形沿斜边的三个三角形块的六种排序．

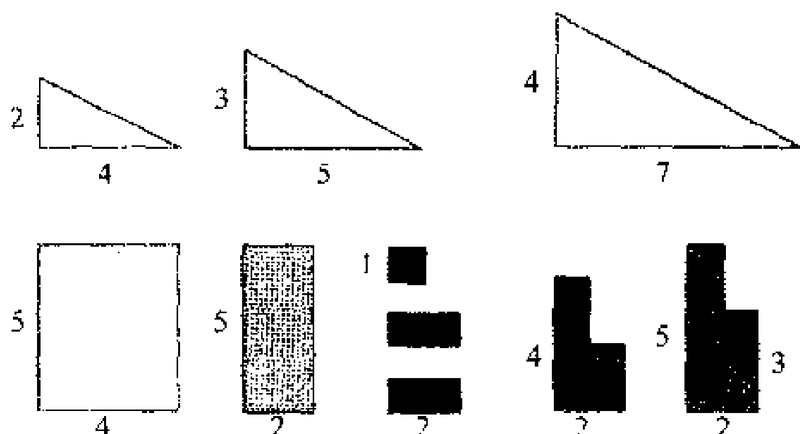


图 16.A 3

对照图 16.A 3 和图 16.A 4,请注意底部右边那个三角形中的非三角形块的总面积为 44,它右边的三角形中那些非三角形块具有总面积 45,下一个是 46,再下一个是 47,再再下一个是 48,而底部左边的三角形是 49. 你可以试着用硬纸板自己制作一套拼块. 让你的朋友们大吃一惊吧!

这里描述的拼图,其中的花招相对来说还是明显的. 然而,布雷特发现了一个构作这种拼图悖论的一般方法,这样就能构作这样的例子:其中沿斜边的三个三角形与那大三角形是如此相似,以致形状上的差异变得根本就察觉不出来.

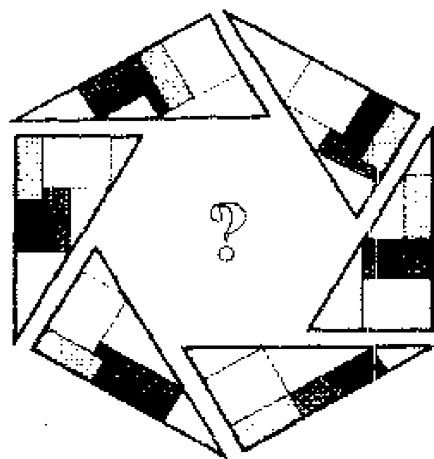


图 16.A 4

第 17 章 填装的三脚架

舍曼·K·斯坦(Sherman K. Stein)

本章专讲一个未解决的问题. 这个问题陈述起来很简单,它甚至可以讲给街上有名的混混儿听. 由于这个问题看来与已知的定理没有关系,任何人,无论是专业工作者还是业余爱好者,同样可以作一尝试. 智力趣题爱好者、计算机程序员或者数学家将发现这是一项诱人的挑战. 况且,它还没有被很多人所研究,因此这是一个产生新方法的好机会.

这个问题说的是把一些整数放到一个正方形阵列的格子里,它是从一个几何问题发展而来的.

对于一个正整数 k,考虑这样一种"三脚架",它由一个单位立方体(作为角顶)和接在它三个不相对的面上的长为 k 的"腿"所组成. 这种 k - 三脚架包含有 $3k+1$ 个单位立方体. 图 17.1 是一个 4 - 三脚架的透视图. 这个问题是这样的:当 k 充分大时,被一个 k - 三脚架通过不重叠的平移放置所能填充的体积在相应的空间范围内占多大比例?

引进一个(x,y,z)坐标系统,其各条轴的正方向对应着这三脚架上三条腿的升出方向. 这个问题就导致了这样一个相关的问题:能有多少个相互不重叠的 k - 三脚架可以把它们的角

[131] 顶立方体放在 $0 \leqslant x,y,z \leqslant k$ 这个边长为 k 的立方体中?

把这些 k - 三脚架的个数称作 $f(k)$. 已知如果当 k 增大时

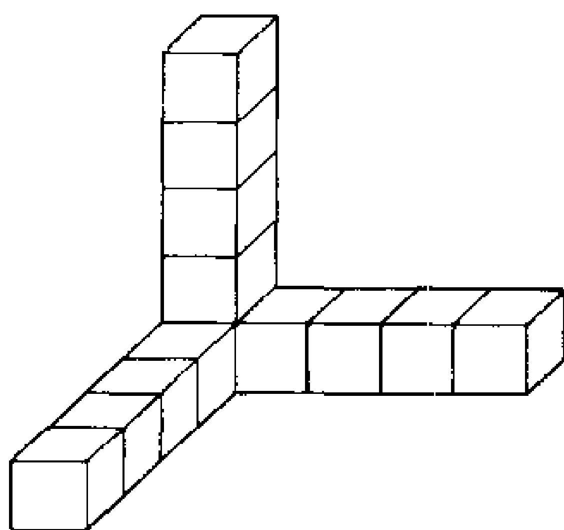


图 17.1

$f(k)/k^2$ 趋近于 0,那么被这些 k - 三脚架所能填装的空间所占的比例也趋近于 0. 因此我们现在面对的问题是:当 k 增大时,$f(k)/k^2$ 趋近于 0 吗?

不难证明,我们可以假设每个角顶立方体与组成那个边长为 k 的立方体的 k^3 个单位立方体中的一个重合. 把每一个这样的单位立方体等同于它那具有最大坐标的顶点,从而也就等同于一个整数三元组(x,y,z),$1 \leqslant x,y,z \leqslant k$. 由于这些三脚架不重叠,所以对于一个给定的数对$(x,y)$,至多只有一个数 z 可以使得三元组(x,y,z)存在. 因此,我们可以在对应于坐标(x,y)的单位格子中填入一个数 z,以这种方式来记录一个三脚架在一种填装方案中的存在. 由于这些三脚架不重叠,所以填好的元素应满足下述"单调矩阵"(monotonic matrix) 定义中的三个条件.

令 k 一个正整数. 一个 **k 阶阵列**(array of order k) 由 k^2 个排成一 $k \times k$ 方阵的空格子组成. 在其中一些格子中以下述三条规则放入 $1,2,\cdots,k$ 中的任何数:

(1) 在每一个(垂直的) 列中,元素从下到上在大小上严格

递增.

(2) 在每一个(水平的)行中,元素从左到右严格递增.

(3) 当我们从左向右看时,被任一个特定的整数所占据的格子一个比一个高("正斜率"条件).

我们称这样一个其中有些格子根据上述三条规则填有数字的阵列为 **k 阶单调矩阵**. 从现在起,我们用单调矩阵代替三脚架填装来作为我们的研究对象.

于是,$f(k)$ 就是所有 k 阶单调矩阵中被占据格子个数的最大值. 显然,$f(k) \leqslant k^2$. 而且由于你可以将数字 $1,2,\cdots,k$ 按序放在单独的一行中,所以 $f(k) \geqslant k$. 现在的问题是,当 k 随意增大时,商 $f(k)/k^2$ 将会发生什么情况? 它有一个极限吗? 这个极限会不会是零?

为了对 $f(k)$ 有个感觉,我们考虑 k 的少数几个小值. 图 17.2 给出了 $1 \leqslant k \leqslant 5$ 的情况. 显然 $f(1) = 1, f(2) = 2$. 花一点点时间就可以证明 $f(3) = 5, f(4) = 8$. $k = 5$ 的情况是由乔伊(K. Joy)编写程序用计算机进行穷尽搜索才解决的,结果证明 $f(5) = 11$. 他找到的许多解答中的一个显示在图 17.2 中. 这些就是仅有的已知 $f(k)$ 的 k 值.

[132]

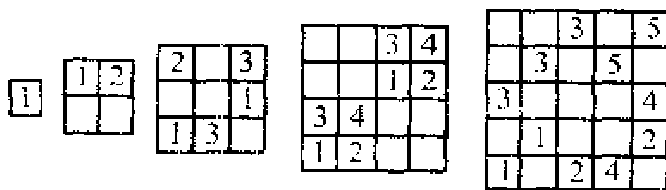


图 17.2

当 k 是一个平方数时,总存在着一个 k 阶单调矩阵,其中有 $k^{3/2}$ 个被占据的格子. 图 17.3 和图 17.2 中第四个阵列说明了这种结构. 这提示我们把 $f(k)$ 写成 $f(k) = k^{e(k)}$ 的形式,并研究 $e(k)$ 的行为. 首先,人们猜想当 k 增大时 $e(k)$ 趋近于 3/2. [这将推出 $f(k)/k^2$ 趋近于 0.] 已知 $e(k)$ 确实趋近于某个数. 显

						7	8	9
						4	5	6
						1	2	3
			7	8	9			
			4	5	6			
			1	2	3			
7	8	9						
4	5	6						
1	2	3						

图 17.3

然,这个数不会大于 2. 称这个数为 L. 结果证明,L 是大于或等于所有 $e(k)$ 的最小数.

知道了这些情况之后,我们考虑图 17.4. 这个单调矩阵证明了 $f(7) \geqslant 19$. 在这个式子的两边取对数,得 $e(7) \geqslant \lg 19/\lg 7 \approx 1.513$. 于是 L 至少是 1.513.

如果你仔细观察图 17.4,你会注意到它的基础是图 17.2 所示的 3 阶单调矩阵,图 17.4 中那九个划分出来的区域的每一个对应着 3 阶阵列中的一个格子. 同样的技巧给出了图 17.5 所示

		4		6		7
	4					5
4				5	7	
				1	2	3
2		3	7			
		1	6			
1	3		5			

图 17.4

			5				8	9
		5					6	7
	5				8	9		
5					6	7		
					1	2	3	4
		3	4	9				
		1	2	8				
3	4			7				
1	2			6				

图 17.5

的 9 阶单调矩阵，其中有 28 个被占据的格子（比图 17.3 中的多一个）. 这表明 $e(9) \geqslant 1.516$，从而 $L \geqslant 1.516$. 藏在图 17.4 和图 17.5 的结构背后的想法证明了

[133]

$$f(2k+1) \geqslant 2f(k) + 3k.$$

令 k 和 l 为正整数. 把一个 kl 阶阵列分割成 l^2 个尺寸为 $k \times k$ 的块，你就能证明

$$f(kl) \geqslant f(k)f(l).$$

同样显然的是 $f(k+1) \geqslant f(k) + 1$. 正是这最后的两个不等式，加上 $f(k) \leqslant k^2$ 这个事实，通过一个已知的定理推出了当 k 增大时 $e(k)$ 有一个极限.

在确定 $f(k)$ 的行为方面，下一步有好几种可能的做法. 其中之一是利用计算机再确定一些 $f(k)$ 值，或至少找出 $e(k)$ 的较大值. 希克森证明了（没有用计算机）$f(255)$ 至少是 4638，这表明 $L \geqslant 1.523$. 不过，可能存在一个小阶数的单调矩阵，它表明 L 还要大.

在还没有人发现当 k 增大时 $f(k)/k^2$ 会发生什么情况的时候，遵循一个光荣的传统，我将提出一个可能相对简单一些的

[134] “练习”问题，它同样没有解决.

令 j 是一个固定的正整数. 令 $g(j,k)$ 是一个 k 阶阵列中能被 1,2,…,j 这些数字以前面所述的三条规则所占据的格子的最大个数. 已知对每个 j,当 k 增大时 $g(k,j)/k$ 趋近于一个极限. 把这个极限记为 $c(j)$.

首先,$g(1,k)=1$.(把一些 1 放在一个 k 阶阵列的上升的斜对角线上.)于是 $c(1)=1$. 对于 $j=2$,就把一些 1 和 2 纳入图 17.6 中所指明的格子. 在这个构造中,每三列有四个元素,这证明 $c(2)$ 至少是4/3. 已知 $c(2)=4/3$. 同样已知的是,$c(3)=5/3$,$c(4)=2$. 于是,对于 $j=1,2,3,4$,$c(j)=(k+2)/3$. 这个模式提示,$c(5)$ 应该是 7/3,但我们只知道它介于 16/7 和 5/2.

						1
					2	
					1	
			1	2		
		2				
		1				
1	2					

图 17.6

顺便说一下,人们可以定义一个在 n 维空间中的 n 脚架. 它由一个角顶立方体和粘在它不相对的面上的 n 支长为 k 的"脚"组成.(当 n 为 2 时,它就像字母 L.)在维数为 1 和 2 的情况下,它可铺砌空间. 正如希克森所指出的,如果 $f(k)/k^2$ 趋近于 0,那么在从 3 开始的所有维数下,n 脚架以一个当 k 增大时趋近于 0 的密度填装 n 维空间. 所以 3 是一个临界维数. 当然,我不知道当 k 增大时 $f(k)/k^2$ 是否趋近于 0. 事实上,我甚至不知道它是否有一个极限.

参 考 文 献

1. W. Hamaker and S. Stein, *IEEE Trans. Inform. Theory* 30(1984), 364—368.

2. S. Stein and S. Szabo, *Algebra and Tiling*, Washington, D. C.: Mathematical Association of America, 1994, Charpter 3.

135

第 18 章 与我的蚂蚁继续同行

戴维·盖尔(David Gale) 吉姆·普罗普(Jim Propp)
斯科特·萨瑟兰(Scott Sutherland)
谢尔盖·特罗别茨科伊(Serge Troubetzkoy)

引 言

本书中一个反复出现的主题是计算机产生的谜团. 例子有用简单的有理递归关系定义的一些序列,结果它们的每一项都是整数,并具有令人感兴感但得不到解释的可除性质;还有实际存在然而没有证明存在的几何构形. 在大多数例子中,所说的谜团一直没有解开,而且对一些情况,可能在一种恰当的意义下确实上是解不开的. 因此,能够对前面所描述的一个谜团提出一个精妙的解答,是一件令人快意的事. 这个解答的一个十分讨人喜欢的特征是,有关的突破是通过画出适当的图形才成为可能的. 这个图形一经画出,什么事情必须予以证明就变得十分清楚. 此后,对这个图形的进一步研究给出了构造这个证明的线索. 到某一步,人们发现原来需要把若尔当曲线定理用到一类特殊的闭曲线上. [137]

在下面这几节中,我们提出一个非正式但(我们希望是)令人信服的论证,它除了要求读者仔细地观察一些图形外,几乎不要求什么了.

迄今为止的故事与有关的谜团

我们再次对某种被称为蚂蚁的自动机产生兴趣. 一个蚂蚁生活在一个被用标准方式划分为一个个方格的平面上,这些方格正如我们所称呼的,叫做**胞腔**(可把胞腔的角顶点看作这平面上的格点). 每个胞腔处于若干种状态中的一种,而且这些胞腔的各种状态随着这胞腔被那个蚂蚁先后访问而一直在发生变化. 在我们准确地解释这蚂蚁与其周围环境是如何相互作用的之前,考虑一个具体的例子将是有益的. 这蚂蚁的一个最简单的非平凡形式最早是由兰顿研究的,其中只有两种状态,我们称之为 L(即"左")和 R(即"右"). 这个蚂蚁最初位于两个胞腔之间的边界上,头朝着东南西北四个基本方向之一(位于垂直的边界上时朝着东或西,位于水平的边界上时朝着北或南). 当这蚂蚁经过它前面那个胞腔时,它要做个 90°转向. 如果那胞腔处于 L状态就向左转,如果处于 R 状态就向右转. 接着便走向相邻胞腔的边界. 当蚂蚁离开一个胞腔时,它就改变这胞腔的状态,把处于 L 状态的胞腔转变为处于 R 状态,或者反之. 对这个蚂蚁的一个非正式的讨论见[1],其中可观察到许多有趣的蚂蚁行为. 就我们现在的目的而言,我们只关注下述仍未得到解释的现象:如果最初所有的胞腔处于同一状态,那么在许多时候这蚂蚁的"踪迹"(它访问过的胞腔的集合,加上这些胞腔当前的状态)呈中心对称,也就是说,由处于 L 状态的胞腔和处于 R 状态的胞腔所组成的构形具有中心对称性. 在图 18.1 中我们复制了相应的图形,其中处于 R 状态的胞腔画成黑色,处于 L 状态的胞腔画成白色. 所有的胞腔最初都处于 L 状态,而蚂蚁开始时位于中央那个胞腔的右边界上,正要向西进入这个胞腔. 这些中心对称性后来停止出现,而且过了大约 10 000 个时间单位,蚂蚁开始表现出一种周期性的"公路建造"行为,向着西南方向的无穷远处而去. 这种"瞬态对称"(transient symmetry)的现象

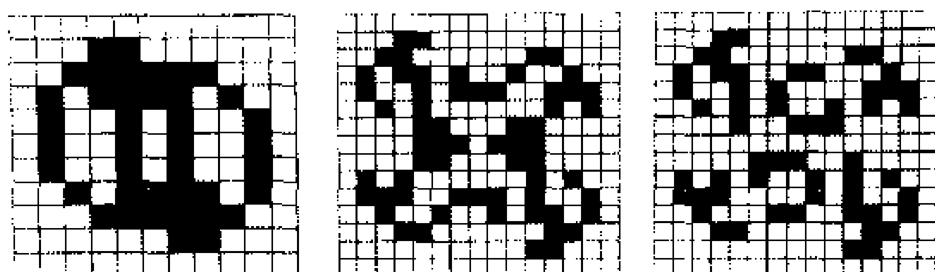


图 18.1 蚂蚁 2 在第 184 步、第 368 步和第 472 步时的宇宙

正等待着人们作出满意的解释.

在[2]中,普罗普描述了更一般的 n 状态蚂蚁的行为. 现在胞腔有着编号为从 1 到 n 的共 n 个不同的状态,而且是由这蚂蚁的内部"程序表"告诉它在选择转弯方向时对哪些状态应作 L 状态处理,对哪些状态应作 R 状态处理. 我们用一个**规则串**来表示这张程序表. 这个规则串由 n 个 L 和 R 排成,其第 k 个字母表示当蚂蚁来到一个处于状态 k 的胞腔时的行动方向. 例如,七状态规则串 LLRRRLR 表示,一个蚂蚁当访问到处于状态 1,2,6 [138] 的胞腔时向左转,而当访问到处于状态 3,4,5,7 的胞腔时则向右转. 当一个蚂蚁来到一个 L 胞腔时(即一个处于 L 状态的胞腔)时,就向左转. 当它来到一个 R 胞腔时,就向右转. 当这个蚂蚁离开一个处于状态 i 的胞腔时,这胞腔便变到状态 $i+1$(以模 n 取剩余). 在这套术语下,那个简单的蚂蚁具有规则串 LR.

出于对称性的考虑,显然只要考虑以 L 起始的规则串就够了. 如果我们在规则串中用 1 代替 L,用 0 代替 R,我们就看到每一个蚂蚁对应着一个用二进制表示的正整数. 于是那个简单的蚂蚁就是蚂蚁 2,而我们那具有"基因组"LLRRRLR 的蚂蚁就是蚂蚁 98. 普罗普发现不同的蚂蚁依据其规则串有着极其不同的行为. 有些看起来完全是混沌的,而其他一些则最终会去建造公路. 唯一能够作出的一般性陈述是下面的

蚁学基本定理(布尼莫维奇 – 特罗别茨科伊) 如果一个

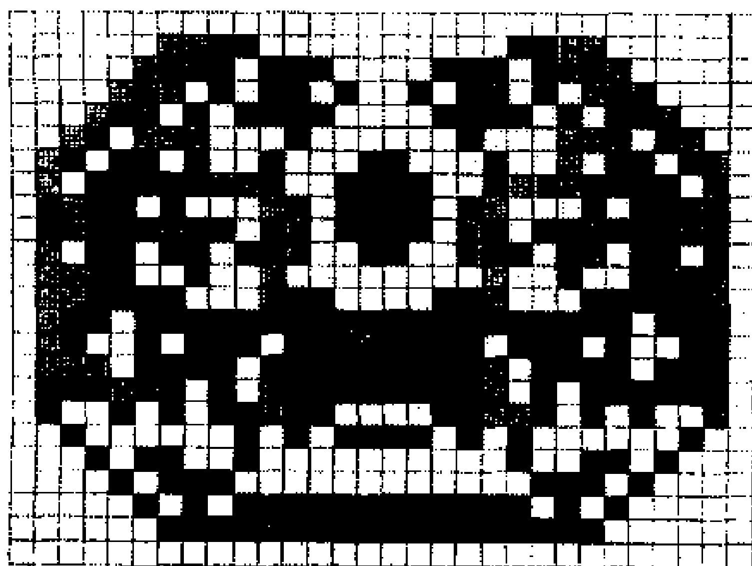
蚂蚁的规则串中至少有一个L和一个R,那么这蚂蚁的踪迹总是无界的.

([1]中有一个证明,可参见.)

在这篇说明性文章中,我们专注于普罗普如下描述的现象:

> 蚂蚁9和蚂蚁12[用我们的符号是LRRL和LLRR]是[在具有长度为4的规则串的蚂蚁中]真正令人惊奇的东西.在这两种情况中,图案虽然越变越大,但总是不会太偏离双侧对称模式!更为特别的是,人们发现蚂蚁频繁地造访它出发时的那个胞腔,而当这种情况发生时,整个构形经常呈现双侧对称.

图18.2和图18.3是从[2]中复制的,它们显示了蚂蚁12在走了16 464步后的踪迹和蚂蚁9在走了38 836步后的踪迹①.这些图形用黑色、白色以及两种灰度来表明四种不同的状态.普



[139] 图18.2　蚂蚁12在第16 464步时的宇宙

① 图18.3与本书第13章的图13.7相同,但那里说这是蚂蚁12在走了186 848步后的情况,而蚂蚁9走了38 836步后的情况如图13.8.——译注

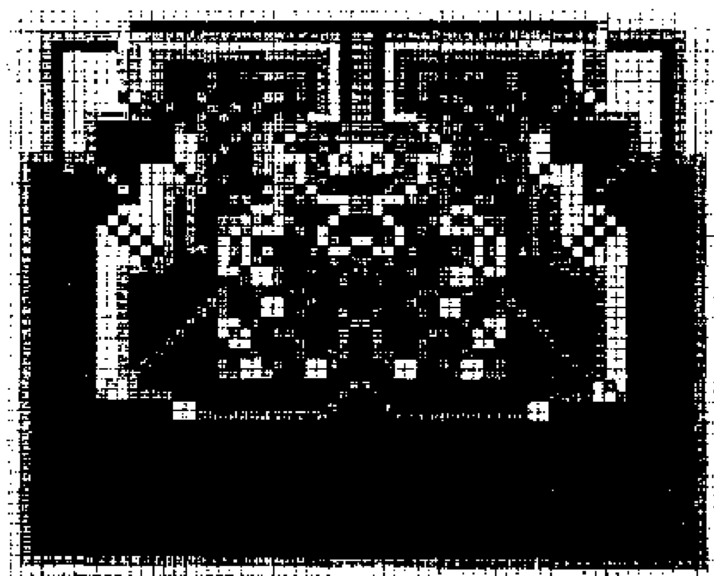


图 18.3　蚂蚁 9 在第 38 836 步时的宇宙

罗普继续写道，“为了再找出这种类型的蚂蚁，我们不得不走进长度为 6 的规则串．在这里，我们遇到了另一个谜团：长度为 6 的规则串中导致双侧对称图案的有 33，39，48，51，57 和 60．注意所有这些数都能被 3 整除！这肯定不是一种巧合．”这的确不是巧合，我们就来证明这一点．

特吕谢[①]铺砖

首先作一个初步的观察．既然蚂蚁每走一步都要转过一个直角，这就推出它的运动是沿水平方向和垂直方向交替着进行的．因此，这平面上的胞腔就像国际象棋棋盘上的格子那样分为两种：一种是 H 胞腔，对它蚂蚁总是从东边或西边沿水平方向进入，而从上边或下边沿垂直方向离开；另一种是 V 胞腔，对它

① 特吕谢（Sébastien Truchet，1657—1729），法国僧侣，专长水力学．在参加奥尔良运河工程时遇到了如何把一种被对角线分为两种颜色的正方形铺砖按不同方位拼合起来作铺砌的问题，因此作了仔细研究，并于 1704 年发表了有关结果．后人称之“特吕谢铺砌系统”．此外，他在印刷字体设计方面也有建树．——译注

蚂蚁总是沿垂直方向进入,沿水平方向离开. 接下来,胞腔的一个至关重要的性质是,它能改变自己的状态,从而改变蚂蚁下次访问了它之后将采取的行动方向(左转或右转). 表示这个性质的关键想法是由吕姆勒提出的,他竟然在这些胞腔中装上了图[140]解式"转换器",这种转换器的方位指明了胞腔当前是处于L状态还是处于R状态. 这些被称为"特吕谢铺砖"的转换器可用图18.4和图18.5来说明. 图18.4显示了一块H铺砖,它先是处于L方位,后是处于R方位;图18.5显示了一块V铺砖的同样情况(一块"H铺砖"是指与一个H胞腔相联系的特吕谢铺砖,V铺砖类似). 请注意,转换方位相当于将这块铺砖作关于其垂直轴或水平轴的镜射.

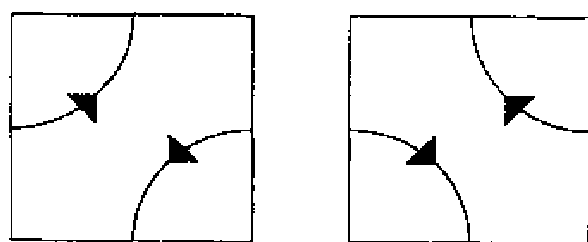


图 18.4　处于L方位和R方位的H胞腔

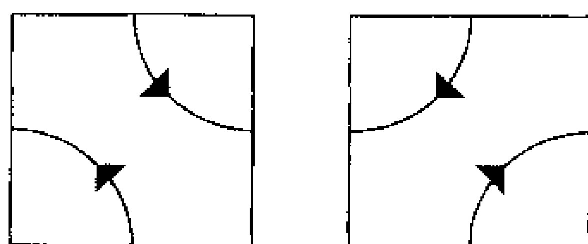


图 18.5　处于L方位和R方位的V胞腔

图18.6显示了用特吕谢铺砖所铺盖的平面,所有这些铺砖都处于L方位. 它给出了由不相交的圆周所构成的一幅图案:"初始构形". 当蚂蚁运动时,其中一些铺砖依据给定的规则串发生转换,但显然这图案将总是由一组不相交的简单闭曲线构成,这些曲线我们称为**特吕谢周线**. (不可能产生无限的周线,

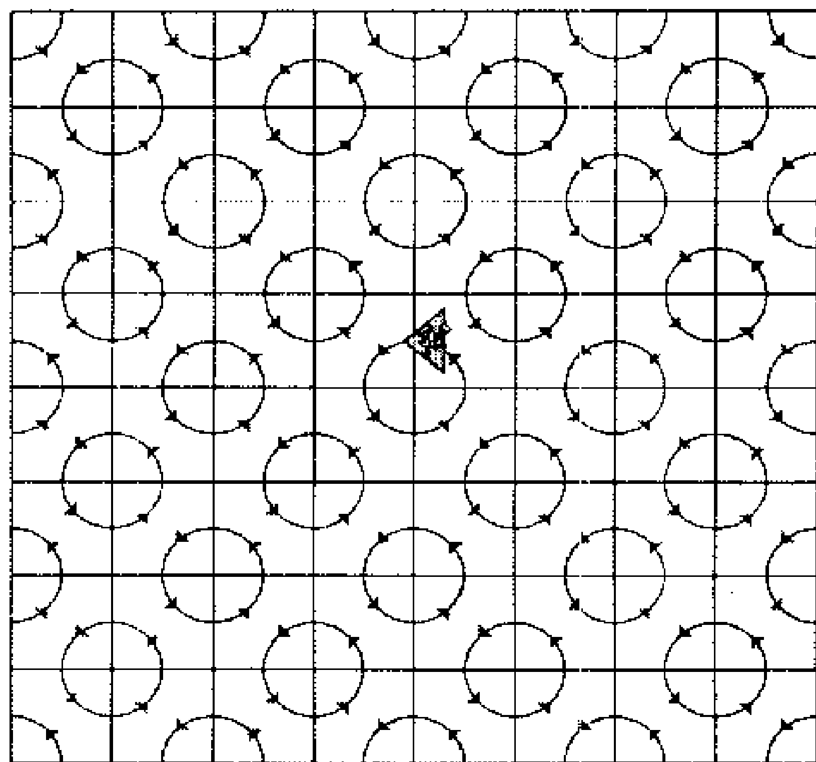


图 18.6 初始构形

因为在每一阶段蚂蚁只转换了有限多块铺砖.）把蚂蚁想像成实际上是沿着特吕谢曲线行进而不是沿着连接胞腔中心的格子道路行进，而且蚂蚁的初始位置是中央那块特吕谢铺砖的某条边中点，那将是有益的（至少目前是如此）. 图 18.7 给出了对应于图 18.1 中那简单蚂蚁的（瞬态）中心对称构形的“特吕谢图 [141]

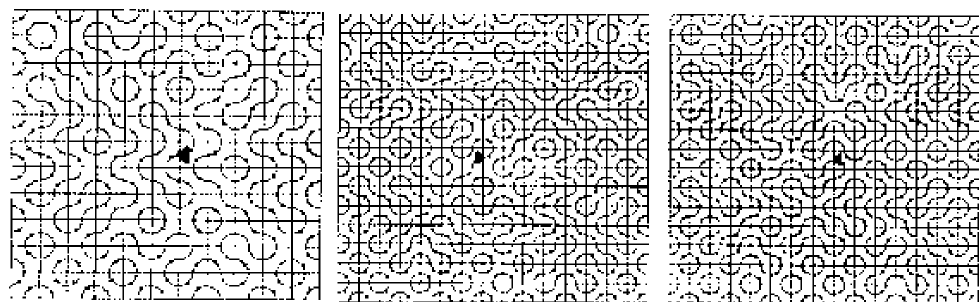


图 18.7 蚂蚁 2 在第 184 步、第 368 步和第 472 步时的宇宙

形",而图18.8和图18.9给出了对应于图18.2和图18.3的特吕谢图形,它们表现了同样的双侧对称性.在所有这些图形中,蚂蚁都回到了它的最初位置,而特别令人感兴趣的正是通过这一点的特吕谢周线,我们称之为**主周线**.开始时,主周线同其他周线一样,只是一个圆周.在图18.9中,对主周线作了强化处理.

[142]

现在,如果人们知道蚂蚁每次从其出发位置离开直到它再次回到这个出发位置都将始终呆在主周线上,那么就可以推出,当蚂蚁完成了它的旅行时这个宇宙在初始状态时的所有对称性都将得到保持,因为每有一个胞腔受到访问,与它对称的胞腔也会受到访问.

不幸的是,一般来说一个蚂蚁不会老是呆在主周线上,因为(如图18.10所示)可能有一些胞腔被这条周线通过两次.这意味着当蚂蚁回到这样的一个胞腔时,这个胞腔的方位可能已经

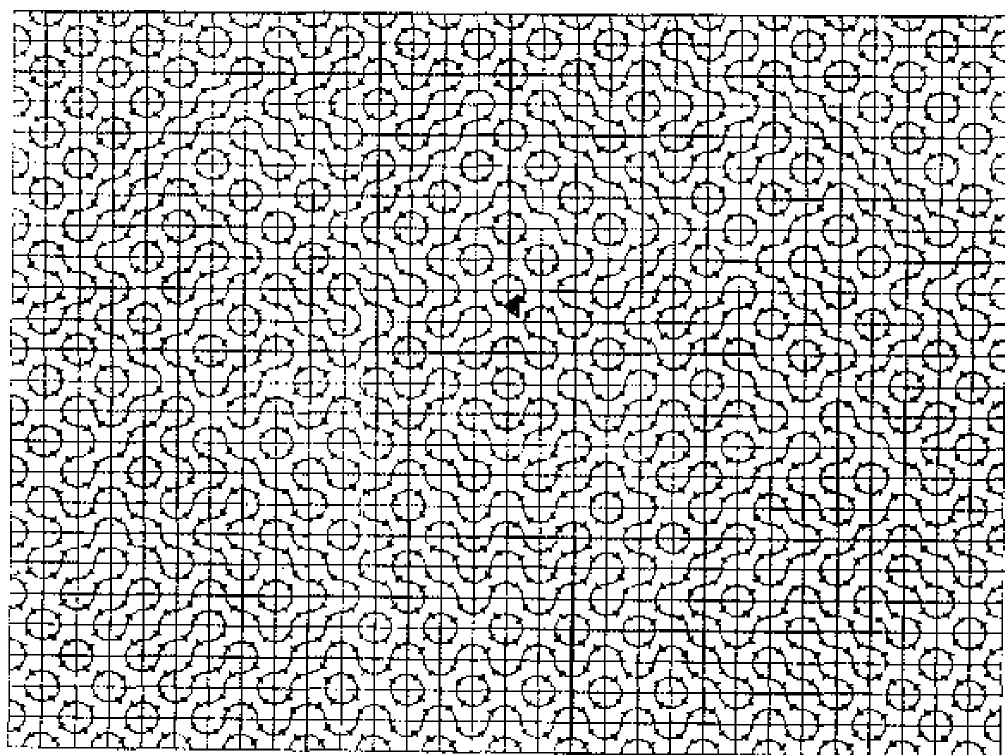


图 18.8 蚂蚁12在第16 464步时的宇宙



图 18.9 蚂蚁 9 在第 38 836 步时的宇宙，对主周线作了强化处理

发生了转换(事实上对兰顿所研究的简单蚂蚁来说，这种情况总是发生)，从而使这蚂蚁脱离主周线. 然而，人们可以证明，事实上对于普罗普所列出的蚂蚁来说，

(1) 被一条周线访问两次的胞腔在它第一次被访问时绝不会发生转换.

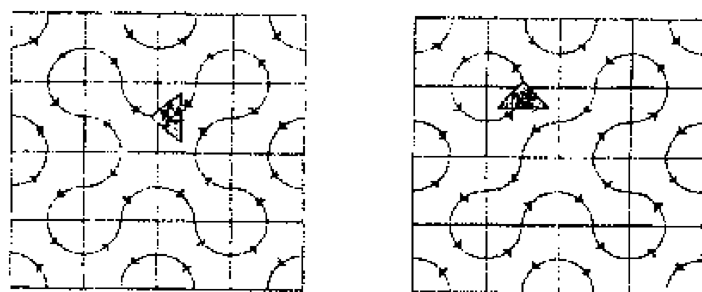


图 18.10 转换:蚂蚁 2 在第 4 步和第 5 步时的宇宙

这表明蚂蚁进行着一种一再地勾画出呈双侧对称的主周线的过程,结果当蚂蚁每次回到其出发点时都造成了一个呈双侧对称的宇宙.性质(1)首先由吕姆勒就蚂蚁12的情况得到证明,后来又被特罗别茨科伊推广到其他蚂蚁上.下面的论证是普罗普对这些证明作了重新加工的结果.

同字母串长度为偶数的性质与增广图

为什么有些蚂蚁的踪迹反复展现出双侧对称性而其他蚂蚁则不?这里是普罗普文章中反复展现出对称性的四状态和六状态蚂蚁的规则串:

蚂蚁	规则串
9	LRRL
12	LLRR
33	LRRRRL
39	LRRLLL
48	LLRRRR
51	LLRRLL
57	LLLRRL
60	LLLLRR

不难看出这些规则串有一个共同之处.如果我们把它们首尾相接,看作是循环排列而不是线性排列,那么它们每一个都是由偶数个L后接偶数个R而构成的.一般地,如果一个规则串在循环排列下是由长度都为偶数的L字母串和R字母串交替排列而形成的,我们就说它具有**同字母串长度为偶数的性质**(even run-length property),例如LRRLLRRRRL.为了使下面的叙述简单明确,我们将只考虑规则串以偶数个L起始的情况(蚂蚁12,48,51和60).对其他情况的论证类似.

为什么这个同字母串长度为偶数的性质能导致双侧对称性

的反复出现呢？在这里，吕姆勒用一种下面将予以解释的方式对有关的图形作了增广．如果一个胞腔的状态是奇数，我们就说它是**冷的**，这种胞腔在下一次被访问时将不会改变方位(由于同字母串长度为偶数的性质)．如果状态是偶数，就说它是**热的**，这种胞腔在下一次被访问时可能改变也可能不改变方位．为了完整地表示这幅图景，我们采用这样一种约定：对于热胞腔，我们不但画出特吕谢铺砖，而且还画出它的对角线．以这些热铺砖的对角线作为边而构成的图，就称为**对角线图**．图18.11和图18.12显示了蚂蚁12和蚂蚁48的对应于它们某几个回到初 [144]

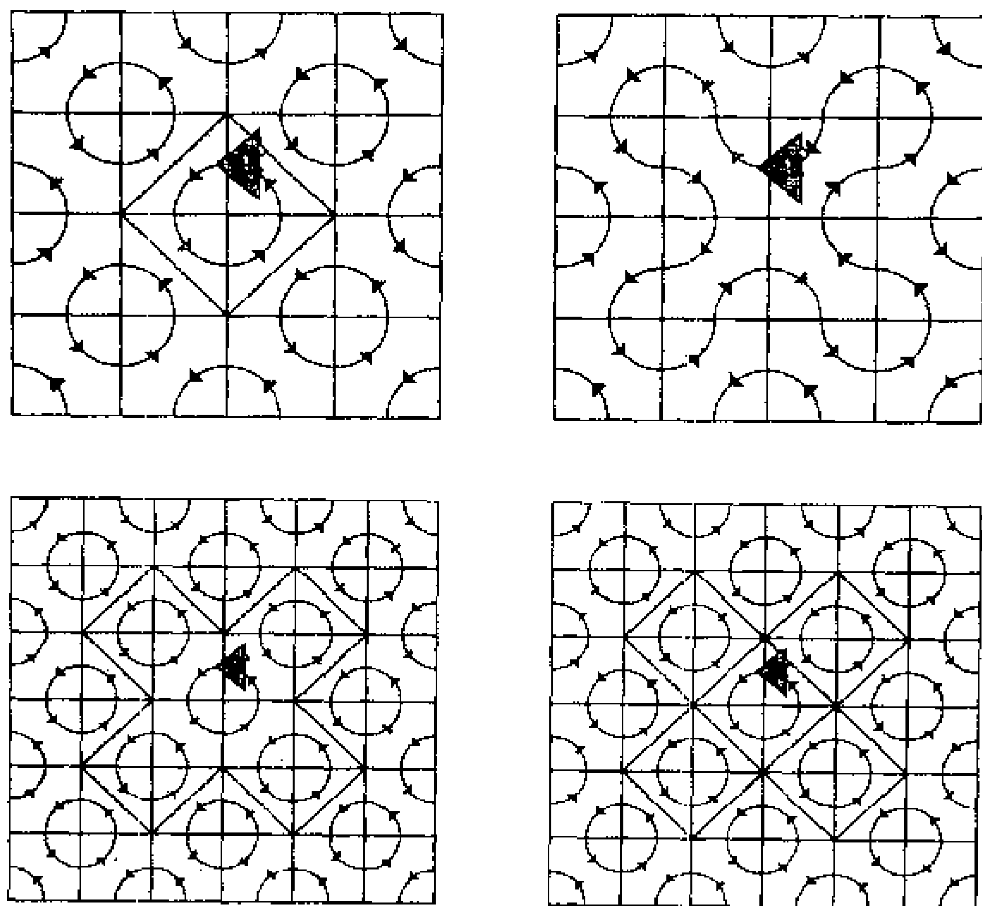


图 18.11 蚂蚁12在它开头四次回家时的对称性宇宙：第4步、第8步、第28步和第32步

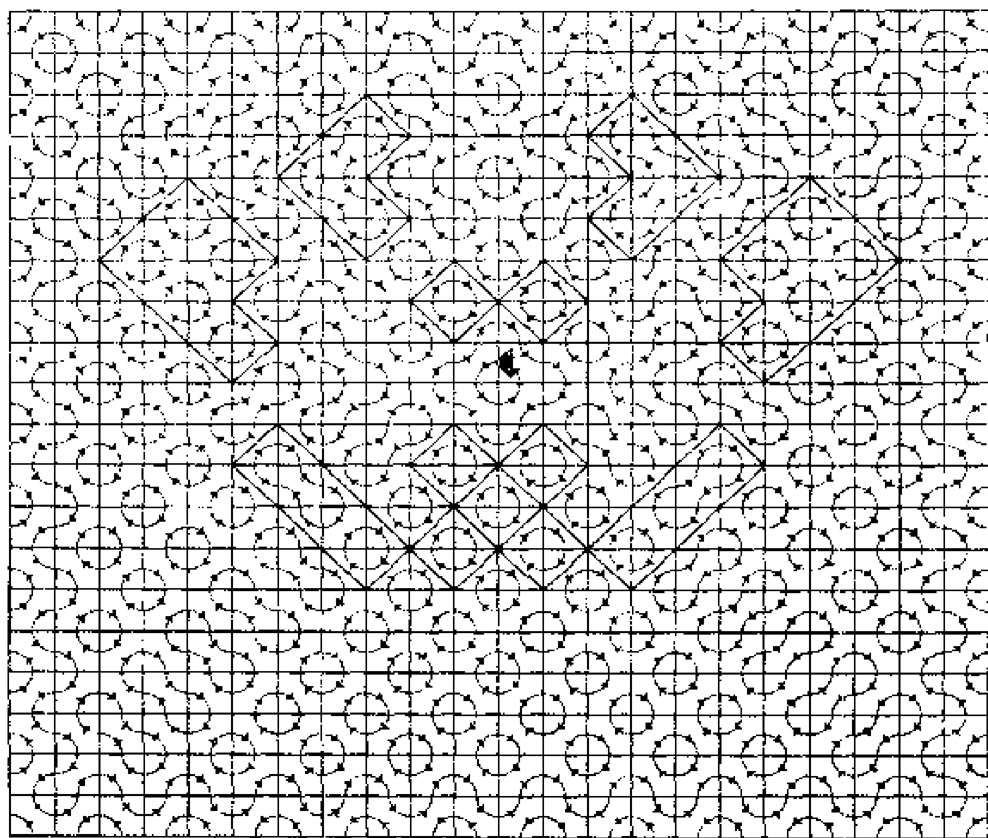


图 18.12　蚂蚁 48 在第 7016 步时的对角线图

始位置的瞬间的对角线图(以后我们将把初始位置称为"家").这些图可能分为许多连通分支,就像图 18.12 所示的那样,十分复杂.然而,请注意一个关键性的事实,我们称之为**对角线次数为偶数的性质**(even diagonals-degree property):

(2) 对角线图中所有顶点的次数[①]都为偶数(例如,次数为 0,2 或 4).

---

① 图论中的概念一般比较直观,不熟悉图论的读者也能理解,故前面均不作注释.但这里的"次数"似有必要说一下.它在许多教科书中称作"度",即一个顶点所关联的边的数目.——译注

引理 1　假设蚂蚁处在家的位置，并且这个宇宙的状态满足(2)，那么这蚂蚁将沿着主周线行进(并回到自己的家).

证明　正如我们先前指出的，我们唯一忧虑的事情是主周线 C 可能对一个胞腔作两次访问，而这个胞腔可能在蚂蚁对它作了第一次访问后改变方位. 如果这个被访问两次的胞腔是冷胞腔，那么相应的特吕谢铺砖在蚂蚁第一次访问后不会改变方位. 如果这被两次访问的胞腔是热胞腔，情况又会怎样呢? 我们将证明，作为(2) 的一个结果，这样的胞腔不存在. 令 T 为这样的一个胞腔，令 d 为 T 中连接顶点 u 和 v 的对角线.

[145]

待证命题　如果把 d 删去，则在所导致的图中，u 所在的连通分支和 v 所在的连通分支不相交.

如果我们能证明这一点，就可导致所期望的矛盾，因为根据(2)，u(或 v) 所在的连通分支将只包含一个次数为奇数的顶点，这就同一个众所周知的事实(通常与欧拉相联系) 发生矛盾：一个连通图必定总是有着偶数个次数为奇数的顶点.（这有时被称为握手定理，它是说握了奇数次手的人的个数是偶数.）

余下的事是证明这命题. 为此，考虑这个被两次访问的铺砖 T 及它的两条弧(四分之一圆周). 不失一般性，我们假定 T 是一块 H 铺砖. 现在把 T 中处于对角线下方的弧涂成红色，另外把 C 中接在这条弧后面直到 C 即将重新进入 T 的那一点为止的所有弧都涂成红色. 把其余的弧涂上蓝色. 图 18.13(a) 和图 18.14(a) 中的虚线代表红色弧，实线代表蓝色弧. 现在考虑同样的图景，只是其中 T 的对角线已被删除而且 T 已被转换成另一方位，如图 18.13(b) 和图 18.14(b) 所示. 与前面一样，T 中处于对角线下方的弧应为红色而其他弧为蓝色①. 人们马上可以

① 图 18.14 中画的是一块 V 铺砖而不是文中假定的 H 铺砖. 相应于 H 铺砖中的对角线“下方”，在 V 铺砖中应是“右方”. —— 译注

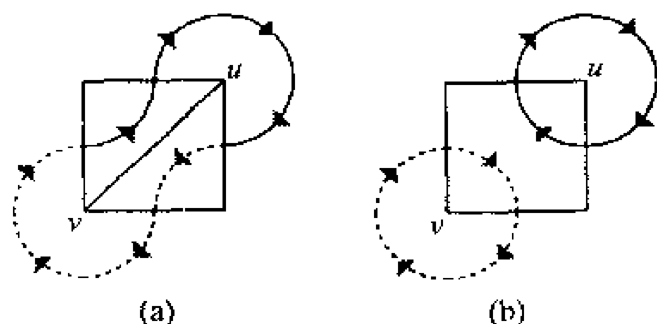


图 18.13 非嵌套情况:(a) 之前;(b) 之后

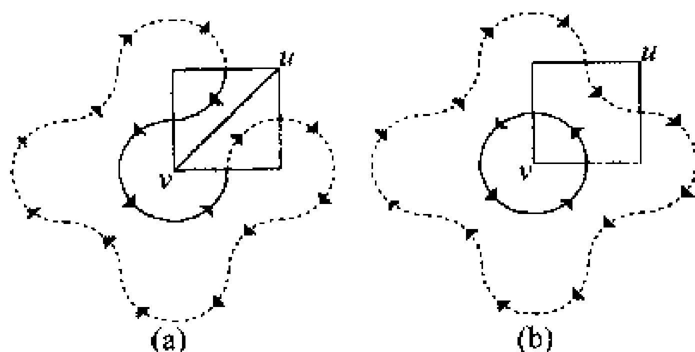


图 18.14 嵌套情况:(a) 之前;(b) 之后

[146] 看到,C 被分裂成两条周线,一条全是红的,另一个全是蓝的(这绝对是一个纯粹的组合数学事件,与这平面的拓扑结构无关).现在若尔当曲线定理告诉我们,这些不相交周线的每一条都有一个内部和一个外部,它们可以像图 18.13(b) 那样布置(非嵌套情况),也可以像图 18.14(b) 那样布置(嵌套情况). 在这两种情况中,u 所在的连通分支和 v 所在的连通分支显然是不相交的,这是因为红色周线和蓝色周线都介于它们之间. 将这个结论与握手定理相结合,即完成了证明.

评注 虽然刚才这些结果不全力使用若尔当曲线定理也可以获得严格证明,但这样必须用到一些平面拓扑,因为在环面上与引理 1 类似的结论不成立.

最后，我们必须证明(2).

引理 2 如果当蚂蚁在家的时候(2)成立，那么当这蚂蚁沿着主周线旅行了一周回到家的时候它仍然成立.

证明 令 v 是对角线图的某个顶点. 以 v 作为一个顶点的四个胞腔组成了**邻居** $N(v)$. 我们证明，如果在某个时候特吕谢周线进入 $N(v)$ 然后离它而去，那么 v 的次数的奇偶性不变. [147]

情况 Ⅰ 这周线只与 $N(v)$ 中的一个胞腔相遇. 那么 v 不在那个胞腔的对角线上. 如果这胞腔是冷胞腔，情况就保持不变，因为冷胞腔不会发生转换. 如果它是热胞腔，那么它就变为冷胞腔，从而就没有了实对角线[①]. 因此，不管它是否发生转换，v 的次数不变.

情况 Ⅱ 这周线与 $N(v)$ 中的不止一个胞腔相遇. 那么令 E 为它进入的胞腔，而 E' 为它离开的胞腔. 如果 E 是热胞腔，那么它的对角线就与 v 关联. 因此当它变为冷胞腔时，这条对角线就消失了(不管 E 是否发生转换). 如果它是冷胞腔，它就变为热胞腔(没有发生转换)，因此将产生一条新的实对角线与 v 关联. 完全同样的证明对 E' 也适用. 因此这两种变化的净效应保持了 v 的次数的奇偶性(见图 18.15). 至于中间经过的胞

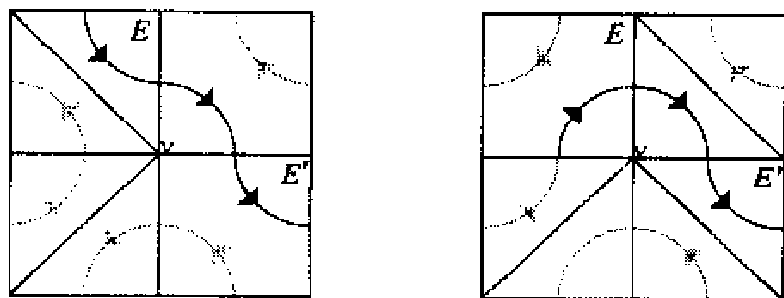


图 18.15 与 v 关联的实对角线的数目其奇偶性不变

① 一个胞腔有两条对角线，在热胞腔中所画的显然是那条不与其中两条弧相交的对角线. 作者现在把这种对角线称为“实对角线”(solid diagonal).—— 译注

腔(它们可能有一个或两个),它们的对角线不与 v 相遇,所以有关的证明与情况 I 相同. 这些胞腔不会对 v 所关联的实对角线的数目造成任何变化. 另外还有一种情况,即 E 与 E' 是同一个胞腔,但我们把有关的分析留给读者.

一般地说,主周线可能出入 $N(v)$ 好几次,但如果人们考察这条曲线在任何进入点与相应离开点之间的部分,前面的分析仍然适用. 这就完成了证明.

结合引理 1 和引理 2,我们终于看到了所发生的情况:如果当蚂蚁在家的时候这宇宙的状态满足对角线次数为偶数的性质,那么这蚂蚁必定是沿着主周线行进,而当它走完这条道路回到家的时候,这对角线次数为偶数的性质又恢复了,因此它必定是再次沿着这条(新的)主周线行进,如此下去,以至无穷. 鉴于上一节末尾的论述,这个定理的证明(以及这个对称之谜的解答)就此完成.

我们把一项任务留给读者:用数论知识简单地证明,一个像 57 这样其二进制展开具有同字母串长度为偶数的性质的正整数能被3整除. 这就解释了普罗普的观察结果,即那些其踪迹反复展现双侧对称的蚂蚁的密码数字.

注记　要复制普罗普的一个程序 ant.c(一个为 UNIX 机设计的蚂蚁宇宙模拟器),请发电子邮件到 propp@math.mit.edu. [148]

参考文献

1. D. Gale, The industrious ant, *Mathematical Intelligencer* 15(2)(1993),54—58(即本书第 10 章).
2. D. Gale and J. Propp, Further ant-ics, *Mathematical Intelligencer* 16(1)(1994),37—42(即本书第 13 章).
3. L. A. Bunimovich and S. Troubetzkoy, Recurrence properties of Lorentz Lattice Gas Cellular Automata, *J. Statist. Phys.* 67(1992),289—302. [149]

[150]

第19章 鞋带问题

约翰·H·霍尔顿(John H. Halton)

你遇到过这样的事吗？你刚刚结束了将纯粹数学中某些美丽的结果耐心地解释给一群非数学家听的努力，希望自己已传达了这颗真美之珠的一些风采，一阵静场之后，有人说，"不错，但这些东西与日常生活有什么关系呢？"我左思右想，认定正确的应答是说"什么也没有．这是对它的过分挑剔．毕竟，日常生活往往是一种拖累，于是我们做数学，其理由就像我们听音乐和上山滑雪一样：脱离日常生活，高于日常生活，超越日常生活"．

但是现在我必须承认，事情偶尔会是另一种样子，日常生活原来是有趣得令人意外的数学的一个来源．这方面的一个美妙例子就是由霍尔顿撰写的这一章[1]．

在许多关于应该怎样系鞋带的讨论中，已经显然没有人似乎有确定的答案．鞋带系了又系，火气越来越大，鞋子甚至被扔了出去……．笔者认定有必要向数学求助．

[151]

这个问题是"流动推销员问题"[2]的一个特例．我们有一个由$2(n+1)$个点(即鞋带孔，或鞋眼)组成的集合，这些点成双行排列，形成一个格点阵列，如图19.1所示．

[1] 这段文字是盖尔的按语．——译注

[2] 关于这个问题，可参见本译丛中《数学：新的黄金时代》的第11章．——译注

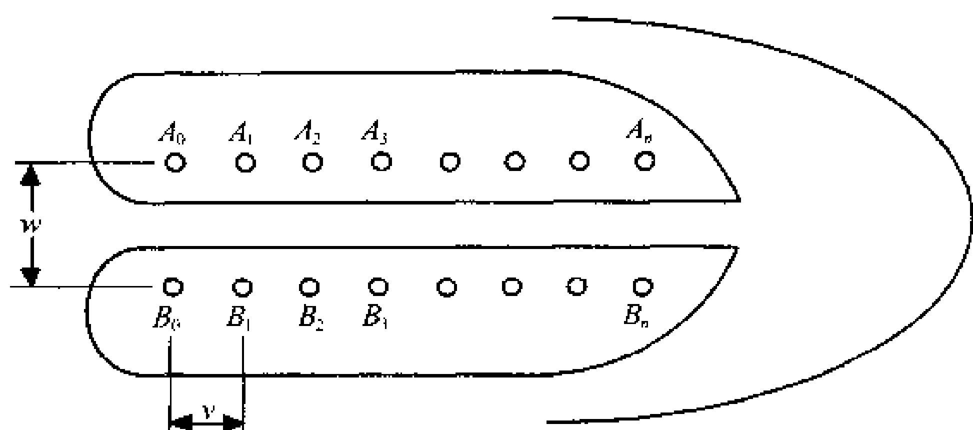


图 19.1 鞋子(一个图解)

这问题是要找出从 A_0 到 B_0 的一条经过每个鞋眼仅一次的最短道路,并且使得子集

$$\begin{aligned} A &= \{A_0, A_1, A_2, \cdots, A_n\} \\ \text{和} \quad B &= \{B_0, B_1, B_2, \cdots, B_n\} \end{aligned} \tag{1}$$

中的点在这条道路上交替出现.

图 19.2 ~ 图 19.4 显示了三种标准的系鞋带策略.

对于美国式(AM),如图 19.2,如果 n 是**奇数**,则系带法是

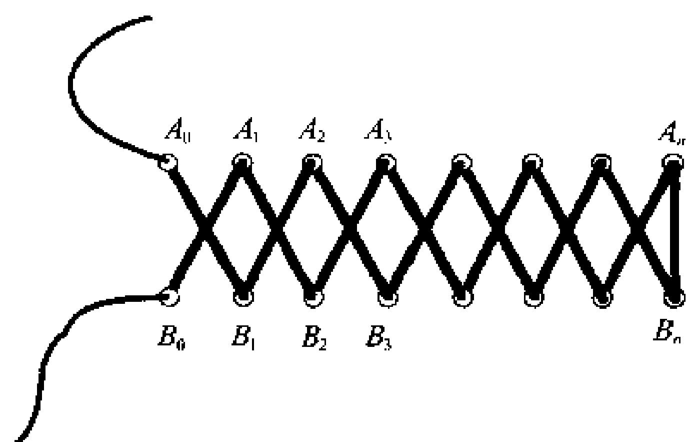


图 19.2 美国式曲折系带法

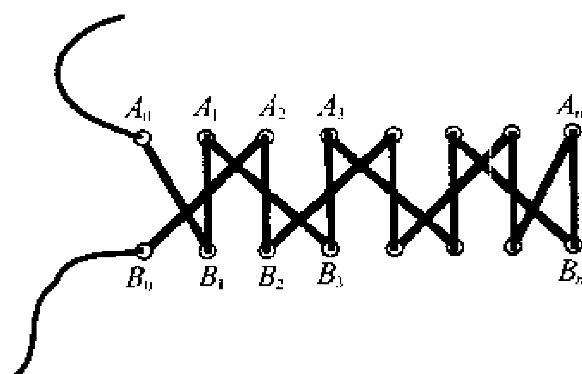


图 19.3 欧洲式平直系带法

$$
\begin{aligned}
&A_0 \to B_1 \to A_2 \to B_3 \to A_4 \to \cdots \\
&\to A_{n-1} \to B_n \to A_n \to B_{n-1} \to A_{n-2} \to B_{n-3} \to \cdots \\
&\to A_3 \to B_2 \to A_1 \to B_0.
\end{aligned} \tag{2}
$$

[152]

如果 n 是偶数,则类似地,其系带法是

$$
\begin{aligned}
&A_0 \to B_1 \to A_2 \to B_3 \to A_4 \to \cdots \\
&\to A_{n-2} \to B_{n-1} \to A_n \to B_n \to A_{n-1} \to B_{n-2} \to \cdots \\
&\to A_3 \to B_2 \to A_1 \to B_0,
\end{aligned} \tag{3}
$$

而且容易验证,在这两种情况中,所用鞋带的总长由下式给出:

$$
L_{\mathrm{AM}} = L_{\mathrm{AM}}(n, v, w) = w + 2n\sqrt{v^2 + w^2}. \tag{4}
$$

对于欧洲式(EU),如图 19.3,当 n 是奇数时,系带法是

$$
\begin{aligned}
&A_0 \to B_1 \to A_1 \to B_3 \to A_3 \to \cdots \\
&\to A_{n-2} \to B_n \to A_n \to B_{n-1} \to A_{n-1} \to B_{n-3} \to \cdots \\
&\to B_2 \to A_2 \to B_0.
\end{aligned} \tag{5}
$$

当 n 是偶数时,类似地,其系带法是

$$
\begin{aligned}
&A_0 \to B_1 \to A_1 \to B_3 \to A_3 \to \cdots \\
&\to A_{n-1} \to B_n \to A_n \to B_{n-2} \to A_{n-2} \to B_{n-4} \to \cdots \\
&\to B_2 \to A_2 \to B_0,
\end{aligned} \tag{6}
$$

而且稍微多想一下,我们就知道在这两种情况中,鞋带总长由下

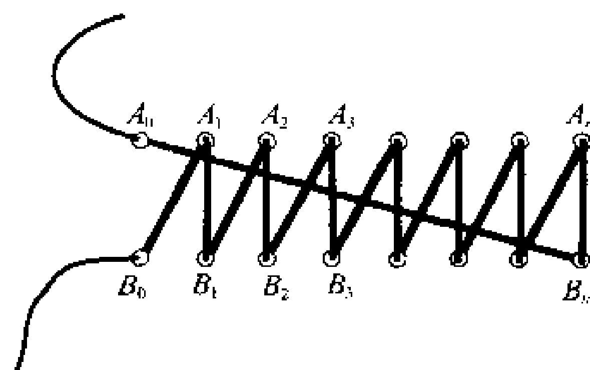


图 19.4 鞋店式快速系带法

式给出：

$$L_{EU}=L_{EU}(n,v,w)$$
$$=nw+2\sqrt{v^2+w^2}+(n-1)\sqrt{4v^2+w^2}. \tag{7}$$

[153]

对于鞋店式(SS)，如图 19.4，系带法是

$$A_0\to B_n\to A_n\to B_{n-1}\to A_{n-1}\to\cdots$$
$$\to B_3\to A_3\to B_2\to A_2\to B_1\to A_1\to B_0, \tag{8}$$

而且我们发现总长由下式给出：

$$L_{SS}=L_{SS}(n,v,w)=nw+n\sqrt{v^2+w^2}+\sqrt{n^2v^2+w^2}. \tag{9}$$

我们可以把这些情形作如下推广．令 α 和 β 表示 $\{1,2,3,\cdots,n\}$ 的排列：

$$\alpha=\{\alpha_1,\alpha_2,\cdots,\alpha_n\},$$
$$\beta=\{\beta_1,\beta_2,\cdots,\beta_n\}. \tag{10}$$

这些排列将由下面的系带法予以对应：

$$A_0\to B_{\beta_1}\to A_{\alpha_1}\to B_{\beta_2}\to A_{\alpha_2}\to B_{\beta_3}\to\cdots$$
$$A_{\alpha_{n-1}}\to B_{\beta_n}\to A_{\alpha_n}\to B_0. \tag{11}$$

它具有总长

$$L=\sqrt{\beta_1^2v^2+w^2}+\sqrt{(\alpha_1-\beta_1)^2v^2+w^2}$$
$$+\sqrt{(\beta_2-\alpha_1)^2v^2+w^2}+\sqrt{(\alpha_2-\beta_2)^2v^2+w^2} \tag{12}$$

$$+ \cdots + \sqrt{(\beta_n - \alpha_{n-1})^2 v^2 + w^2} + \sqrt{\alpha_n^2 v^2 + w^2}.$$

对于上面所示的那三种具体的系带法,有关的特定排列为

$$\begin{aligned} \alpha_{AM} &= \{递增的所有偶数后接递减的所有奇数\}, \\ \beta_{AM} &= \{递增的所有奇数后接递减的所有偶数\}; \end{aligned} \tag{13}$$

$$\alpha_{EU} = \beta_{EU} = \beta_{AM}; \tag{14}$$

$$\alpha_{SS} = \{递减的所有数\}. \tag{15}$$

这些排列简单得确实令人惊奇.

定理 1 如果 $v = 0$ 或 $w = 0$,则对所有正的 n,

$$L_{AM} = L_{EU} = L_{SS}. \tag{16}$$

如果 $v \geqslant 0$ 并且 $w \geqslant 0$,则

$$L_{AM}(1, v, w) = L_{EU}(1, v, w) = L_{SS}(1, v, w), \tag{17}$$

而如果 $v > 0$ 并且 $w > 0$,则

$$L_{AM}(2, v, w) < L_{EU}(2, v, w) = L_{SS}(2, v, w). \tag{18}$$

最后,如果 $v > 0, w > 0$,并且 $n > 2$,则

$$L_{AM} < L_{EU} < L_{SS}. \tag{19}$$

[154]

这个定理可用(4),(7)和(9),通过仔细分析情况并消去根号而得到证明. 这证明留给读者作为一个练习.(它已由笔者在一份技术报告[1]中给出.)

格点表示

让我们把集合 A 和集合 B 平行等距地交替放置,从而形成一个如图 19.5 所示的格点阵列. 给出任何一个系带法 $\mathscr{L}$,我们都能它表示成一条折线(分段的直线段)L,就像图 19.5 中对我们那三种标准系带法例子所表示的那样. 这条折线总是向下伸展,在这新的格点阵列中穿行,对每个鞋眼仅访问一次.

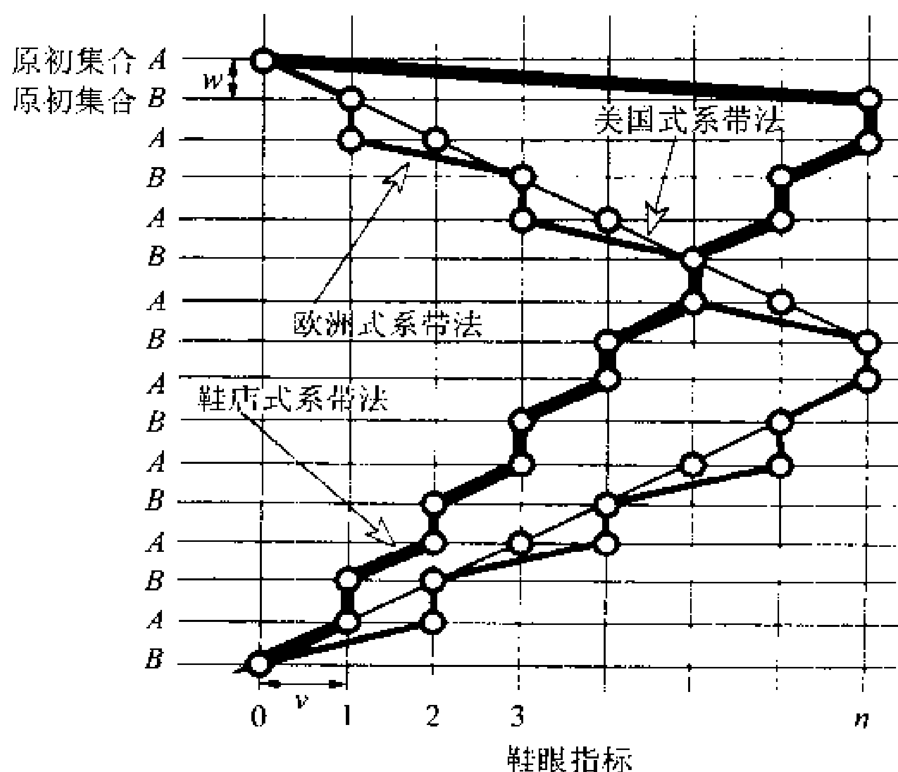


图 19.5 三种标准系带法的格点表示

按系带次序的第一条线段 $A_0 \rightarrow B_{\beta_1}$ 其位置不变；接下来的一条线段 $B_{\beta_1} \rightarrow A_{\alpha_1}$，被它关于原初那条 B 线的镜象所替代；再下一条是 $A_{\alpha_1} \rightarrow B_{\beta_2}$，它被平行下移了两个格点间距（也可以说这是一个被相继作了两次镜射而得到的象），如此等等；而最后一条线段 $A_{\alpha_n} \rightarrow B_0$，则回到了把原初那条 B 线上的 B_0 向下作 $2n$ 个格点间距的位移而得到的象．显然，L 这个折线表示的总长 [155] 等于相应系带法 $\mathscr{S}$ 原来的总长 L.

现在，直接应用一下三角不等式，美国式系带法的总长短于欧洲式系带法这一事实立即就变得十分显然（见图 19.6）.

L_{AM} 和 L_{EU} 这两个表示在几个地方发生重合．在它们不重合的地方，出现了一个三角形 QPR 及其一系列的复制品，显然有 $PR < PQ + QR$，(19) 中的第一个不等式因此推出，不需要再

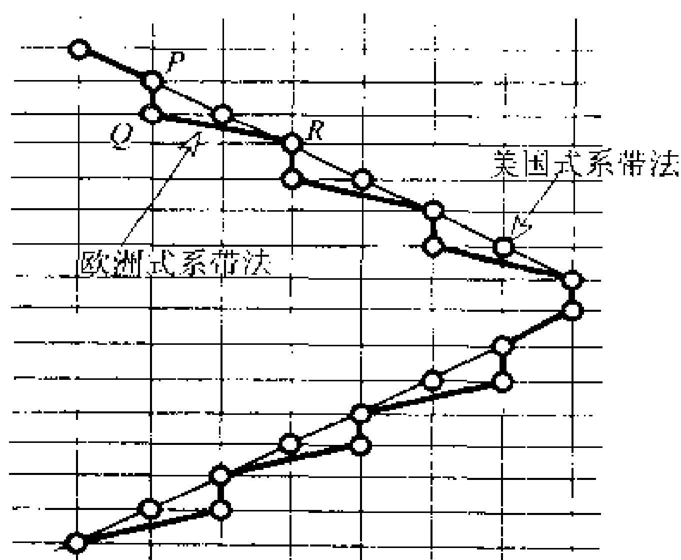


图 19.6 美国式系带法与欧洲式系带法的比较

用到代数了! 156

欧洲式系带法优于鞋店式系带法这一事实的证明稍微困难一些(见图 19.7). 首先,我们观察到表示 L_{EU} 和 L_{SS} 只有**两条**对角线是相同的,它们都是在两个方向上各行进一个格点间距(斜

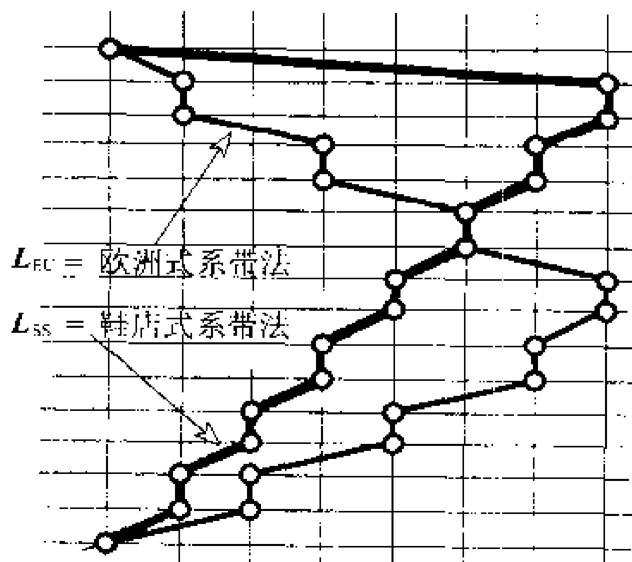


图 19.7 欧洲式系带法与鞋店式系带法的比较

率为 $\pm w/v$)；相同的还有 n 条(垂直)线段，它们都是仅在垂直方向上行进一个格点间距 w. 如果我们删去所有这些相同的线段，把下面被分离的线段向上平行移动(在前两种情况中，还要作侧向移动)，与上面的线段重新会合，这相当于**从每个表示中减去相等的长度**，我们就得到了简化表示 L_{EU}^{*} 和 L_{SS}^{*}. 这个结果显示在图 19.8 中. 现在每个表示都由一条只折了一次的折线构成(仅有两条相接的线段 —— 一条这个方向，一条那个方向).

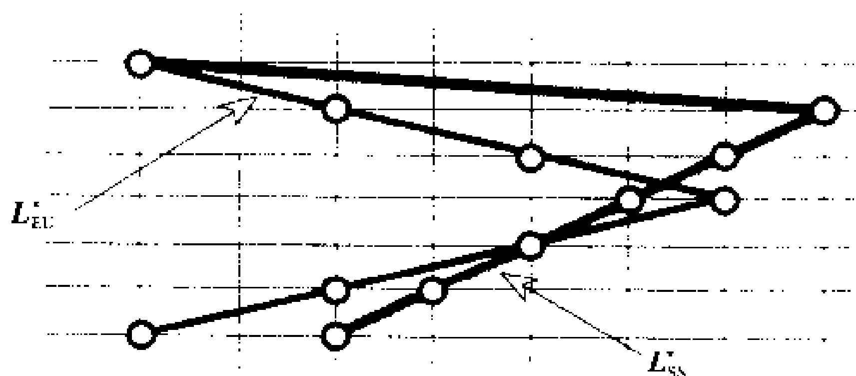


图 19.8　欧洲式系带法与鞋店式系带法的比较 —— 简化表示

现在我们再次表演那个"镜射技巧"，这次是沿水平坐标方向，即把每个表示中那条向左的线段作关于垂直轴的镜射. 这样得到的表示线记为 L_{EU}^{**} 和 L_{SS}^{**}(见图 19.9).

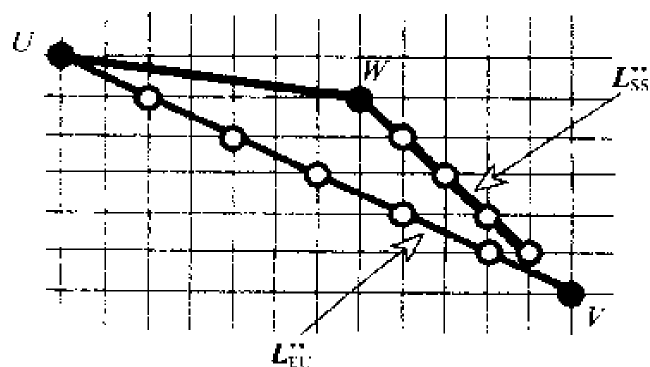


图 19.9　欧洲式系带法与鞋店式系带法的比较 —— 镜射表示

[157]

现在我们可以直接观察到，L_{EU}^{**} 只是一条单一的直线段 UV，而 L_{SS}^{**} 由两条直线段 UW 和 WV 构成，因此还是由三角不等式，(19) 显然成立.

最　优　化

我们采用上面描述的格点表示(见图 19.5— 图 19.7)，并将"镜射技巧"应用于道路上从 B_n 到 B_0 的部分. 对应于一种典型的一般系带法的道路其样子示于图 19.10. 其中还显示了对应于美国式系带法的道路 L_{AM}^{**}. 在这个特定的例子中，同前面一样，令 $n=7$，而那一般的系带法是

$$
\begin{aligned}
&A_0 \to B_2 \to A_7 \to B_4 \to A_6 \to B_1 \to A_1 \to B_3 \to A_3 \\
&\quad B_6 \to A_5 \to B_5 \to A_4 \to B_7 \to A_2 \to B_0.
\end{aligned}
\tag{20}
$$

其长度由下式给出[对照(12) 并归并相同的根式]：

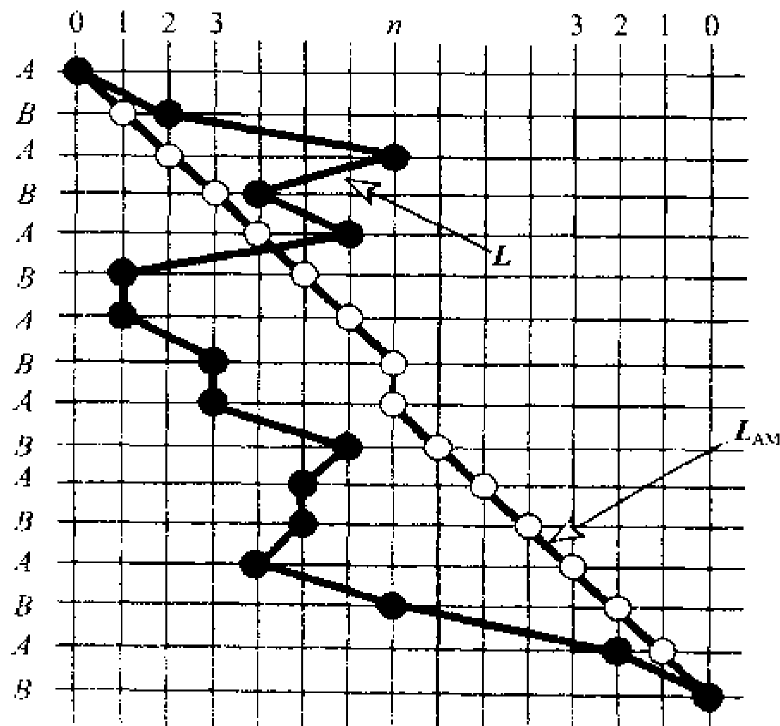


图 19.10　一般系带法——镜射表示

$$L = 3w + 2\sqrt{v^2 + w^2} + 4\sqrt{4v^2 + w^2}$$
$$+3\sqrt{9v^2 + w^2} + 3\sqrt{25v^2 + w^2}. \tag{21}$$

一般地,令系带法具有总长

[158]
$$L = \sum_{k=-n}^{n} = N_k\sqrt{k^2v^2 + w^2}, \tag{22}$$

其中,显然有:

$$\sum_{k=-n}^{n} N_k = 2n + 1 \tag{23}$$

是向下位移的净总数(即步数,因为每一步都有一个距离为一个格点间距 w 的向下位移),而

$$\sum_{k=-n}^{n} kN_k = 2n \tag{24}$$

是以一个格点间距 v 为单位的向右位移的净总数. 对于美国式系带法,显然有

$$N_0 = 1, N_1 = 2n, \quad \text{所有其他的 } N_k = 0. \tag{25}$$

定理 2 美国式系带法具有尽可能短的总长 L,而且它是唯一最优的系带法.

证明 令 $\boldsymbol{L}$ 是任意一种系带法 $\mathscr{L}$的镜射表示,并令 L 是它的总长.

(i) 如果 $N_0 \geqslant 1$,我们就从 $\mathscr{L}$ 中任意拿走一个相应的(垂直)步,并从 $\mathscr{L}_{AM}$ 中拿走那个单独的垂直步,把表示中被分离的两段像前面那样通过平行移位重新会合. 于是这两个新的表示 $\boldsymbol{L}^{+}$ 和 $\boldsymbol{L}_{AM}^{+}$ 仍然具有共同的端点,两者的长度都比原来只减小 w. 现在,L_{AM}^{+} 显然是最小的,因为它是连接这两个端点的直线段. 因此,对所有的 $\mathscr{L}$,

$$L_{AM} \leqslant L. \tag{26}$$

[159] (ii) 现在假设 $N_0 = 0$. 这种情况由图 19.11 表示.

不可能只对 k 的正值有 $N_k > 0$. 如果是这样,则根据(23)

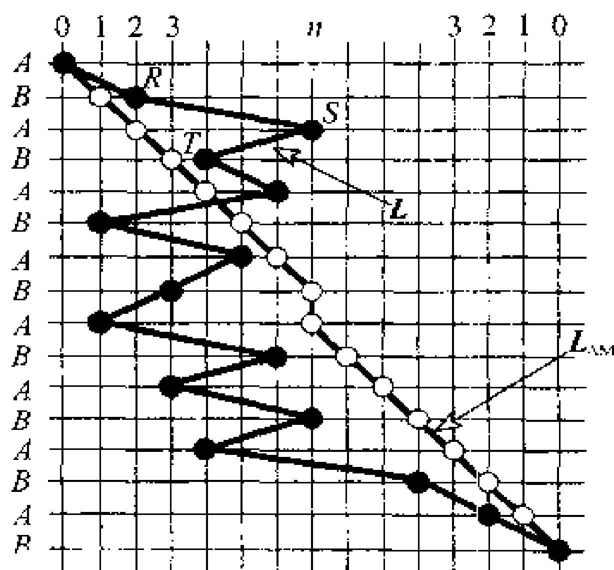


图 19.11 $N_0 = 0$ 的情况 —— 没有垂直线段

和(24),我们将会有

$$\sum_{k=1}^{n} kN_k - \sum_{k=1}^{n} N_k = N_2 + 2N_3 + \cdots + (n-1)N_n = 1. \tag{27}$$

这是不可能的,因为所有的 $N_k \geqslant 0$. 因此,至少有一步具有负的(即向左的)水平位移,从而按从上到下的次序,就存在着一个第一个出现的向左步 ST. 显然这一步**不可能**是这个表示中的**第一步或最后一步**. 于是,它前面是一个**向右**步 RS,这样就形成了一个指向右方的角.

现在(见图 19.12 中的局部放大)令 F 和 G 分别为过 R 和 T 的垂直线与过 S 的水平线相交处的格点. 于是有

$$|FR| = |GT| = w \tag{28}$$

和

$$\begin{aligned} |FS| &\geqslant v, \\ |GS| &\geqslant v. \end{aligned} \tag{29}$$

过 G 画一条直线平行于 TS,并令它与 RS 交于 X(因为它一定会与 RS 相交). 再过 X 画一条垂直线(平行于 RF)与 TS 交于 Y.

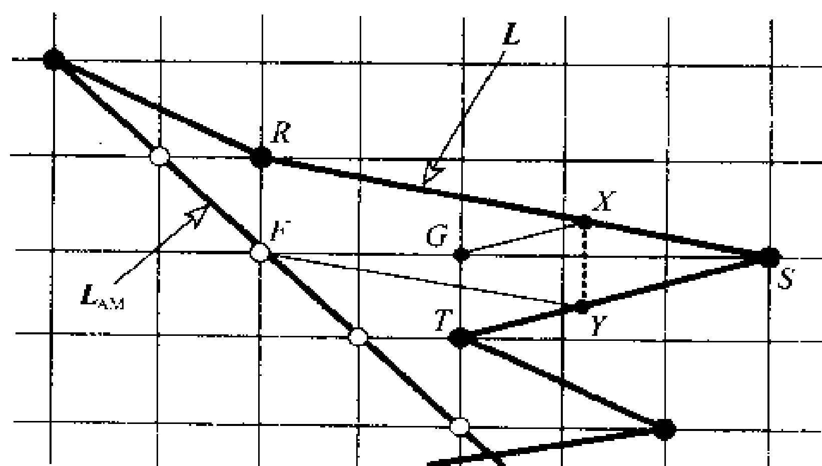


图 19.12　图 19.11 的局部放大

显然 $XYTG$ 是一个平行四边形，因此由(28)，$|XY| = |GT| = w$①.

这样我们就可以用折线 $RXYT$ 来代替表示 L 中的折线 RST，并由三角不等式，

[160]

$$|XY| < |XS| + |SY|, \tag{30}$$

使得这个变形表示 $L^{\perp}$（不妨用这个记号）比 L 短. 但是现在 L^{-} 有一条长度为 w 的垂直线段，所以由与情况（ⅰ）中同样的证明，不等式(26)得以确立.

注意：表示折线 L^{-} 通常**不是任何系带法的一个表示**，因为它一般来说不是由格点连接而成的. 但这没有什么关系，因为证明到这一步，我们只对线的**长度**感兴趣.

我们现在已经证明，如果 $\mathscr{L}_{\min}$ 是任何具有最小长度的系带法，那么它和它的（水平镜射）表示 $L_{\min}$ 将具有与美国式系带法相等的总长，即根据(4)，

$$L_{\min} = L_{AM} = w + 2n\sqrt{v^2 + w^2}. \tag{31}$$

---

① 注意 $XYFR$ 也是一个四行四边形，因为对边 XY 与 RF 平行相等. —— 原注

（iii）最后，我们证明这最优系带法 $\mathscr{L}_{\min}$ 的**唯一性**．在情况（i）和情况（ii）中给出的论述，证明了任何最小系带法 $\mathscr{L}_{\min}$ 都满足(25)；也就是说，它的(水平镜射）表示 $L_{\min}$ 具有 $2n$ 条沿着对角线向右和向下各走过一个格点间距的直线段和一条垂直线段．然而，如(31)所表明的，这条垂直线段在这条链中的**位置**并不对总长 $L_{\min}$ 产生影响．

不过，既然 $L_{\min}$ 不仅是格点折线，而且是一种**系带法**的表示，那么它**一定要**经过那条对应于指标 n 的垂直格线正好两次(即穿过相应的鞋眼 A_n 和 B_n)，而这条线是**仅有的**一条**没有**被镜射变换所复制的格线，因为它**就是**镜射轴．因此，既然这个表示单调向右行进(也就是说，从不向左)，那么那条独一无二的垂直线段就被严格地限制在指标 n 的位置上，就像在 L_{AM} 中那样．这就完成了定理2的证明．

参考文献

1. John H. Halton, The shoelace problem, *Department of Computer Science Technical Report* No. 92-032, University of North Carolina at Chapel Hill, 1992.

[161]

[162]

第20章 三角形与证明

这是我们关于三角形这个话题的第三章，但是它很有些与众不同．前面那两章讨论的是计算机对三角形几何的作用，而我们在这里将返本还源，对一些非常经典的论题进行一种新的考察．

“作出一个推广，其目的并不仅仅在于把更多的情况包括进来，而且还在于把不必要的假设丢弃掉，有时候这可能导致更简单的证明．”我记得在做研究生的时候，我最早的导师之一斯廷罗德向我讲过像这样的一些话．事实上有两种类型的推广．人们既可以减弱假设，也可以加强结论．下面讨论的例子对这两种途径都作了阐明．在所有情况中，找到了正确的推广，就把一个看起来复杂的问题简化为一种本质上机械的验证，即一件只不过是“摇摇把柄”的事情．我们的第一个例子是纽曼给出的，为此我们再次向他表示感激．

莫利[1]的奇观

唐纳德·J·纽曼(Donald J. Newman)

关于当前数学教育的指导思想，一件令人悲哀的事情是对[163]平面几何的回避．如今的这一代，或许还有他们的父母，都没有

[1] 莫利(Frank Morley，1860～1937)，在英国出生的美国数学家．以下面将要介绍的莫利定理而闻名．——译注

听说过像九点圆、笛卡儿定理[1]、切瓦(Ceva)定理[2]这样的奇观，或者那奇观中的奇观，即莫利三角形.

如图20.1所示，我们取一个任意的三角形，并把它的角三等分，从而得到三个交点. 这些点在这个初始的三角形中构成了一个小三角形. 然而，这个内部的小三角形完全不是任意的. 莫利的伟大发现(1899)就是，它总是等边三角形！

当年我阅读，或更准确地说，试图阅读莫利关于这个惊人的定理的证明时，我发觉它是绝对的难以理解. 我对自己说，或许在将来的年岁里我会回来重新阅读，然后把它读懂. 但我在这上面从未获得成功，甚至当我读到那个基于三角学的简单得多的证明，还有纳望辛加(M. T. Navansiengar)给出的那个相当简单的几何证明时，感到其中仍然有着太多的复杂之处，缺乏让人读

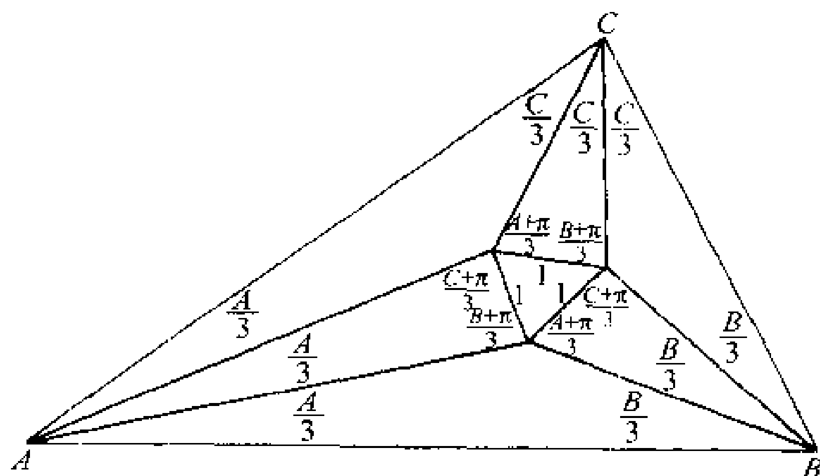


图 20.1

① 虽然笛卡儿的名声如雷贯耳，但笛卡儿定理看来很少有人知道. 设一个立体角的平面角之和为 α，定义 $2\pi-\alpha$ 为其"亏角"(dificiency). 笛卡儿定理是说，一个多面体的顶角的亏角之和为 4π. 可以证明，笛卡儿定理等价于多面体的欧拉公式(关于欧拉公式，可参见本译丛中《数学娱乐问题》的第4章). ——译注

② 关于切瓦定理，可参见本译丛中《近代欧氏几何学》的第8章，但那里译成"塞瓦定理". ——译注

下去的动力.(一连串的机会!)难道我们要永远放弃对这莫利奇观的理解?抑或我们失败是因为我们要求太低?不管怎样,莫利定理还是宣称,在图 20.1 中,那个内部的三角形将永远是等边三角形.所有这些证明之所以看来如此困难而缺乏诱人的动因,可能是因为莫利定理其实只是这故事的一半.全部的图景就在图 20.1 中,它讲出了完整的故事,实际上给出了自我证明!(这种情况在归纳法证明中经常发生:较完全的陈述比有限度的陈述容易证明.)

于是我们求助于[4]中所用的"欺骗"策略,具体地说,我们从那个等边三角形开始向外构作.结果得到图 20.2.在这里我们作了正规化处理,即把那个等边三角形的边长取为 1.注意根据对称性,只要证明那个被标示的角为 $A/3$ 就够了.

[164]

到这一步,我们可以把这个证明转交给一名在三角学课程中学过"解三角形"的高中生去完成了.我们看到,图 20.2 中所有的边长都由那三个构作出的三角形的角 — 边 — 角(ASA)所确定.于是,将正弦定理用于三角形 $AB'C'$ 即得到

$$\frac{AC'}{\sin(C+\pi)/3}=\frac{1}{\sin(A/3)},$$

故

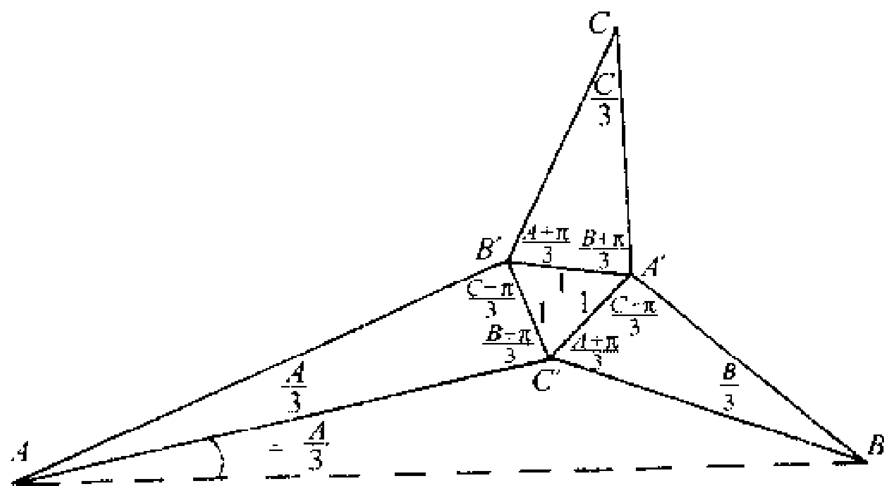


图 20.2

$$AC' = \frac{\sin(C+\pi)/3}{\sin(A/3)}.$$

类似地，

$$BC' = \frac{\sin(C+\pi)/3}{\sin(B/3)}.$$

还有

$$\angle AC'B = 2\pi - \frac{A+\pi}{3} - \frac{B+\pi}{3} - \frac{\pi}{3},$$

因此对三角形 $AC'B$，我们有图 20.3.

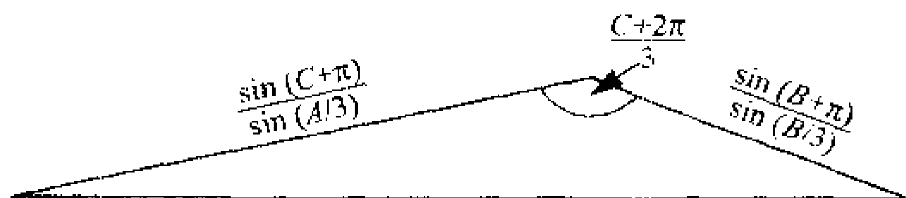


图 20.3

对这个三角形，我们知道了两条边及它们的夹角，(SAS)，所以其余的角都被确定，它们必定是 $A/3$ 和 $B/3$. 对此人们可以再次使用正弦定理予以验证. QED.

参 考 文 献

1. Frank Morley, Extensions of Clifford's theorem, *Amer. J. Math* 51 (1929), 465—472.
2. J.M.Child, A proof of Morley's theorem, *Math. Gaz*. 171(1922).
3. M.T.Navansiengar, *Educ. Times*, *New Series* 15, 47(1909).
4. H.S.M.Coxeter, *Introduction to Geometry*, Toronto: John Wiley & Sons, 1982, 23—25.

一个三角形定理的剖析与演化

正如每一名中学生曾经知道的那样，一个三角形的中线交于一点（形心），高线也是这样（垂心）.（这年头，你能找到一个

居然知道什么是中线的孩子,那就很幸运了.)

关于共点性,一个人们不太熟悉的例子是费马点. 这是(在一个锐角三角形中)与三个顶点的距离之和为最小的点. 为找出这个点,我们以已给三角形的三条边为底边,作出三个等边三角形. 然后把其中每个三角形的远端顶点同那已给三角形中与其相对的顶点连接起来. 这三条连线的交点就是我们要找的点(图 20.4).

更不怎么著名的是下面这个或者说这对有关一个作图的定理,这个定理有时归功于拿破仑. 在图 20.4 中,把 A',B',C' 取为那三个三角形的中心而不是其远端顶点. 连线 AA',BB',CC' 仍然共点.(所谓的拿破仑定理说,在这种情况下点 A',B',C' 本身是一个等边三角形的顶点!)而且,这三个等边三角形可以取在原来那个三角形的外面,也可以取在里面.

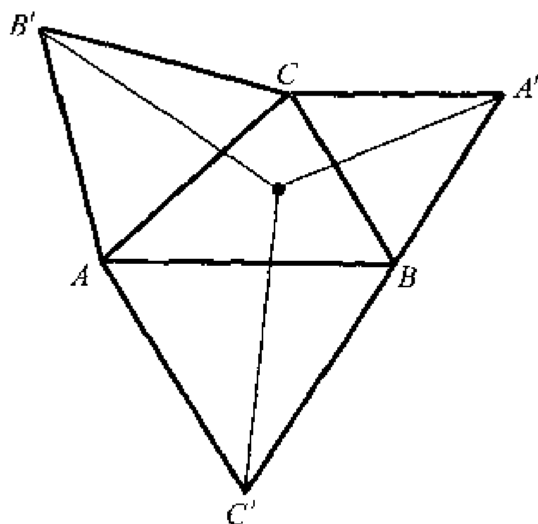


图 20.4　费马点

第一个推广

这些定理是一个无穷的单参数定理族的特殊情况. 这个定理族最早是在 100 多年前被发现的,虽然看起来它在不断地被

重新发现. 这个归功于基佩特[1]的定理把费马和拿破仑的三个等边三角形代之以任意三个互为相似的等腰三角形. 设这些三角形的底角为 α,则费马点对应着 $\alpha = \pi/3$ 的情况,拿破仑点对应着 $\alpha = \pi/6$(正或负)的情况;而形心和垂心分别是极限情况 $\alpha = 0$ 和 $\alpha = \pi/2$(请验证). 当 α 从 $-\pi/2$ 变到 $\pi/2$ 时,那公共交点的轨迹就是一条等轴双曲线,又称基帕特双曲线. 一篇近期的说明性文章见于《数学杂志》(*Mathematics Magazine*)1994 年 6 月号第 288 ~ 205 页.

习题(令人意外但很容易) 证明基佩特双曲线还经过那三个顶点 A, B, C;因此它是经过形心、垂心和这三个顶点共五个点的唯一圆锥曲线. (做一下. 进去玩一把. 你敢吗?)

第二个推广

上面的单参数族原来是一个无穷的三参数定理族的一个特殊情况. 请参见下页的图 20.5,我们有

定理 1 给出一个三角形 ABC 及点 A', B', C',使得 $\alpha = \angle BAC' = \angle B'AC, \beta = \angle ABC' = \angle A'BC, \gamma = \angle A'CB = \angle ACB'$,则 AA', BB', CC' 共点.

于是,三角形 ABC 边上的三角形不一定是等腰三角形. 关键的条件是这三个三角形的底角要像图 20.5 中所示的那样两两相等. 请注意,由于 α, β, γ 是任意的,我们就有了一个三参数的定理族;而基佩特的结果是 $\alpha = \beta = \gamma$ 这种特殊情况. 这个定理陈述起来是那么简单和自然,以致人们怀疑它一定在很久之 [167] 前就被注意到了,然而它的历史踪迹看来朦胧不清. 近期的一篇参考文章是柯比(D.Kirby)的一篇"课堂笔记"(Classroom

[1] 基佩特(Ludwig Kiepert, 1846 ~ 1934),德国数学家. 以下面说到的基佩特双曲线而闻名. —— 译注

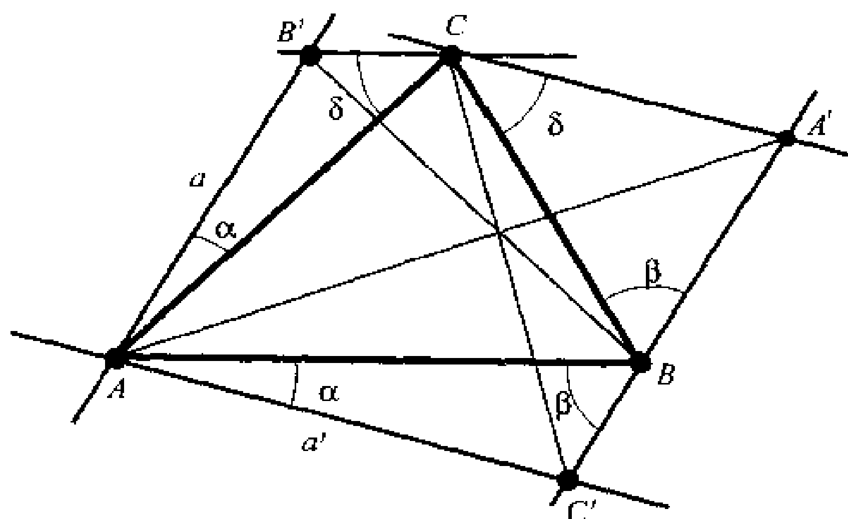


图 20.5

Note,载《美国数学月刊》January(1980),p.45—p.47).

第三个推广

余下的故事基于克利福德·加德纳寄来的一些考察结果.

首先是一些术语:图 20.5 中的直线 a 和 a' 称为关于顶点 A 的**等角线**(isogonal line),意思是它们关于过 A 点的角平分线互为镜象.类似地,我们可以不作关于角平分线的镜射,而是作关于这三角形中线的"镜射".更准确地说,如果过一个顶点的两条直线截对边的交点与这条对边的中点等距,则称它们为**等截线**(isotomic line).如果过顶点的直线是是等截线而不是等角线,则与定理 1 类似的定理同样成立.事实上,这个结论对**任何**三条分别过 A, B, C 的共点直线以下面的形式成立.

定理 2 令 p, q 和 r 分别为过三角形 ABC 三个顶点 A, B 和 C 的共点直线.令 P_A 为过点 A 的线束.令 T_A 为 P_A 上满足下述条件的(唯一的)射影映射:

(1) 令 AB 和 AC 互换;

(2) 令 p 保持不动.

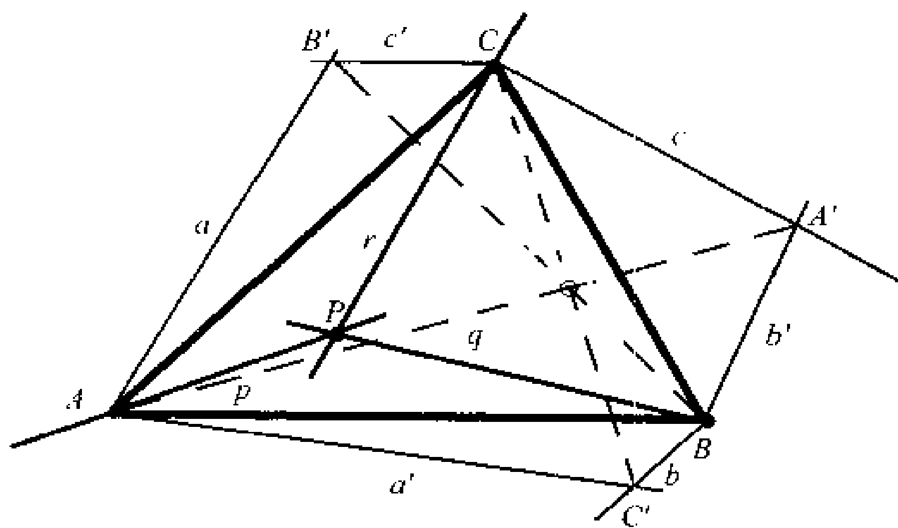


图 20.6

类似地定义 P_B, P_C 和 $T_B、T_C$. 对 P_A 中的任意一条直线 a，令 $a' = T_A(a)$. 类似地，对 P_B 中的 b 和 P_C 中的 c，令 $b' = T_B(b), c' = T_C(c)$. 令 $C' = a' \cap b, A' = b' \cap c, B' = c' \cap a$. 那么，$AA', BB'$ 和 CC' 共点①.

注意现在我们有了一个无穷的五参数定理族，因为点 P 的坐标可以任意选取②.

实际上，定理 2 是定理 1 的一个直接结论，只要我们利用一个众所周知的事实，即平面上的一个射影变换可通过任何四个独立的点予以任意地定义. 作为一个特殊例子，令顶点 $A、B、C$ 保持不动，而把内点变换到任意点 P，图 20.5 就变成了图 20.6. [168] 因此，这个推广的效果并不在于包容更多的情况，而在于简化证

① 这里以及下面证明中关于射影几何的一些概念，可参见《射影几何趣谈》(冯克勤著，上海教育出版社，1987 年版). 此外，原文中对 A' 和 B' 的定义有误，现已作更正. —— 译注

② 点 P 即 $p、q、r$ 的公共交点，而一个射影映射由四个参数决定，故共有五个可变参数. —— 译注

明. 具体地说,我们再构作一个射影变换,它把 C 变到原点,把 A 变到 y 轴上的无穷远点,把 B 变到 x 轴上的无穷远点,而把 P 变到点(1,1),如图 20.7 所示.

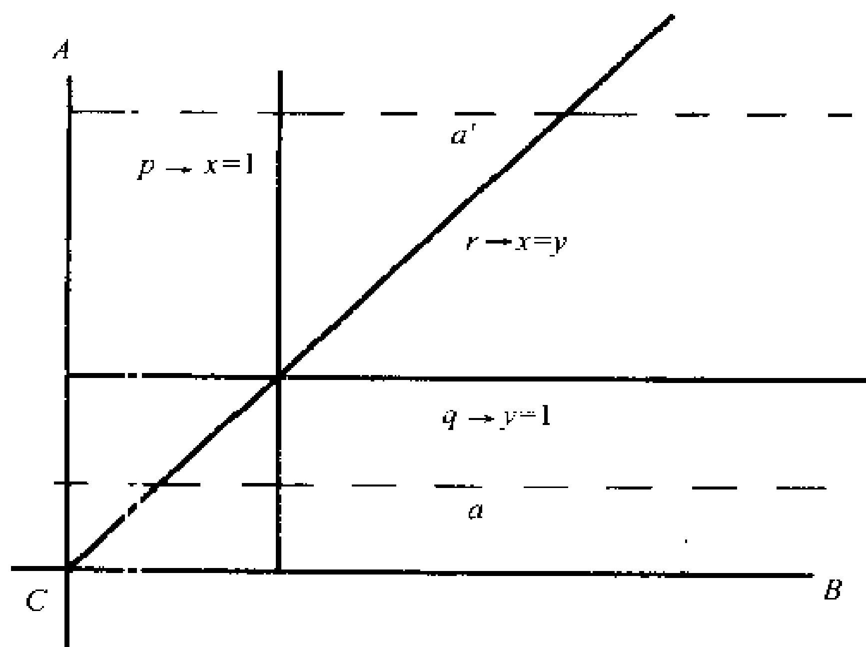


图 20.7

于是直线 p,q,r 就分别变成了直线 $x = 1, y = 1, x = y$,而且我们有

P_A 是所有垂直线的集合,

P_B 是所有水平线的集合,

P_C 是所有过原点的直线的集合.

现在请回忆,任何一维的射影变换都取 $y = (ax + b)/(cx + d)$①的形式,因此 T_A 把直线 $x = a$ 映成直线 $x = 1/a$. 这是把 y 轴与无穷远线互换,并令直线 $x = 1$ 保持不动的的唯一射影变

① 符号使用又混淆了. 这里的 a,b,c,d 是射影变换的参数,x 和 y 是同一线束中元素的一维(非齐次)射影坐标. ——译注

换. 我们将把方程为 $x = a$ 的垂直线简记为 a,把水平线 $y = b$ 记为 b,而把过原点的直线 $y = cx$ 记为 c. 然后根据 T_B 和 T_C 的类似定义,我们有

$$a' = T_A(a) = \frac{1}{a}, b' = T_B(b) = \frac{1}{b}, c' = T_C(c) = \frac{1}{c}.$$

现在,计算点 A', B', C' 的坐标以及直线 AA', BB', CC' 的方程,并验证它们共点,就成了高中解析几何中一件简单的事了. 具体地说,

$$C' = a' \cap b = (1/a, b),$$
$$A' = b' \cap c = (1/bc, 1/b),$$
$$B' = c' \cap a = (a, a/c),$$

于是,AA' 具有方程 $x = 1/bc$, BB' 具有方程 $y = a/c$, CC' 具有方程 $y = (ab)x$. 那么,$AA' \cap BB' = (1/bc, a/c)$,它就在直线 CC' 上. [169]

第四个(最后一个)推广

请注意定理 2 是一个射影几何定理. 这就意味着定理 1 是"绝对的",也就是说它除了在欧氏几何中成立外,在椭圆几何和双曲几何中也同样成立(在椭圆几何的情况下,人们要假设点 A', B', C' 的存在). 即使对最简单的特殊情况,即作为这篇文章出发点的中线共点,这个结论也不是显然的. 常见的欧氏证明需要大大借助于平行公设,因为其中必须让连接一个三角形两边中点的直线平行于第三条边. 使用其他方法可以证明这个中线定理在非欧几何的情况下同样成立. 与此相反,定理2则直接覆盖了所有这三种几何的情况. [170]

第21章 多联骨牌

所罗门·W·戈隆布(Solomom W. Golomb)

希尔伯特第十八问题[①]

1900年,在巴黎国际数学家大会上,当伟大的德国数学家希尔伯特在一篇演说中提出他为20世纪数学所安排的议事日程时,他那份包括着23个未解决问题的“通缉要犯名单”中,有一个编号为18的问题.这是一个关于如何让 n 维欧氏空间(包括 $n = 2$ 和 $n = 3$ 的情况)能用单单一个几何图形的全等副本所“铺盖”(pave)或“铺砌”(tile)的问题.他具体地问道:

> #1.“在 n 维欧氏空间中是不是……只存在有限多种不同类型的具有[紧]基本域的运动群?”[②]
>
> #2.“是不是还存在这样的多面体,它们并不作为运动群的基本域出现,然而以它们为基础,通过一种

---

① 这一章在编辑上似十分奇怪,整章只作一节,即“用多联骨牌铺砌矩形”,且作者姓名署在这个节标题下面.现根据内容分节并拟了相应的标题.——译注

② 这里所说的运动群即全等变换群.所谓一个运动群 Γ 的基本域,粗略地说,就是欧氏空间 X 中的这样一个区域 F,它满足:(1) $\Gamma(F) = X$,即 F 在 Γ 中全等变换的作用下将覆盖整个空间 X;(2) 对于 Γ 中的任何两个不相同的全等变换 γ_1 和 γ_2,$\gamma_1(F)$ 与 $\gamma_2(F)$ 不在它们的内部相交,即(1)中所描述的覆盖是不重叠的.有兴趣的读者可参见《直观几何(上册)》(D.希尔伯特、S.康福森著,王联芳译、江泽涵校订,高等教育出版社,1959年版).——译注

将其全等副本作适当毗连的方式,能把整个[欧氏]空间完全填满?"

#1中所说的运动群被比勃巴赫所确定,但是对#2作出肯定回答的例子却在3维(赖因哈特(K. Reinhardt),1928)和2维(黑施(Heesch),1935)的情况中找到了.更简单、更一般的相关例子后来被鲁滨逊、斯坦以及其他人找到.图21.1显示了一个与黑施的例子相关的例子.(还存在更小的例子,但是要证明它们具有所规定的性质多少比较困难.)不难证明,能让这个图形铺砌平面的唯一方式如图21.2所示.然而,它并不是一个"基本区域",因为唯一能把A移到B的运动是由一个关于那条虚线的镜射和一个向上2个单位的平移所组成的,但是B在这个运动下的象却不是这铺砌样式中的一块铺砖.

希尔伯特根本没预料到会有哥德尔不完全性定理,更不要说他的一些问题会被证明是"不可判定的".就同他那个关于找出任何已给("丢番图(Diophantine)")方程的所有整数解的第十问题被(朱莉娅·鲁滨逊(Julia Robinson)、马季亚谢维奇(Ю. B. Матиясевич)等人)证明是计算不可判定的(computationally undecidable)一样,对于任意给出的一个有限的铺砖集合,是否能用其中铺砖的全等副本来铺砌平面这个一般性问题也被王浩证明

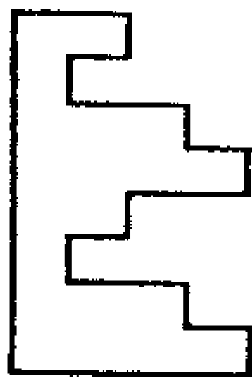


图21.1 一个能铺砌平面但不是一个"基本域"的图形

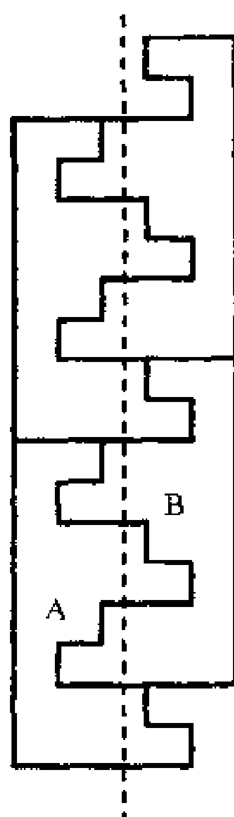


图 21.2 用图 21.1 中的图形铺砌平面

是不可判定的.(王浩一方面证明了这个铺砌问题与图灵机"停机问题"——一个标准的不可判定问题——是等价的;另一方面证明了它的不可判定性与一类包含三个量词的逻辑公式——所谓∀∃∀公式类——的不可判定性是等价的[①].)伯杰(R. Berger)和王浩的其他学生推广了这个结果,例如,他们证明了,是否能用单单一个几何图形的全等副本来铺砌平面(或者一个较小的区域,如一个矩形)的问题也是计算不可判定的.

本章对二维铺砌问题中一特殊情况的已知结果作一概述.

---

① 关于图灵机停机问题,可参见本译丛中《20 世纪数学的五大指导理论》的第 4 章. 关于王浩的结果,可参见《数理逻辑通俗讲话》(王浩著,科学出版社,1981 年版).——译注

用多联骨牌铺砌矩形

我们考虑的铺砖将只限于**多联骨牌**,这里,一个"n 联骨牌"是指取 n 个相同的单位正方形并把它们沿共同边界拼接起来而得到的任何连通图形. 于是,不考虑方位,就只有一种**单骨牌**(即这个单位正方形本身)和一种**2联骨牌**(即多米诺骨牌,这个家族全体成员的名称就是由这位祖先的名称得来的①). 有2种不同的**3联骨牌**,5种不同的**4联骨牌**,12种**5联骨牌**,35种**6联骨牌**,等等. 其中比较简单的示于图21.3. (我们不把互为镜象("镜射象")的多联骨牌看作两种不同的形状.)

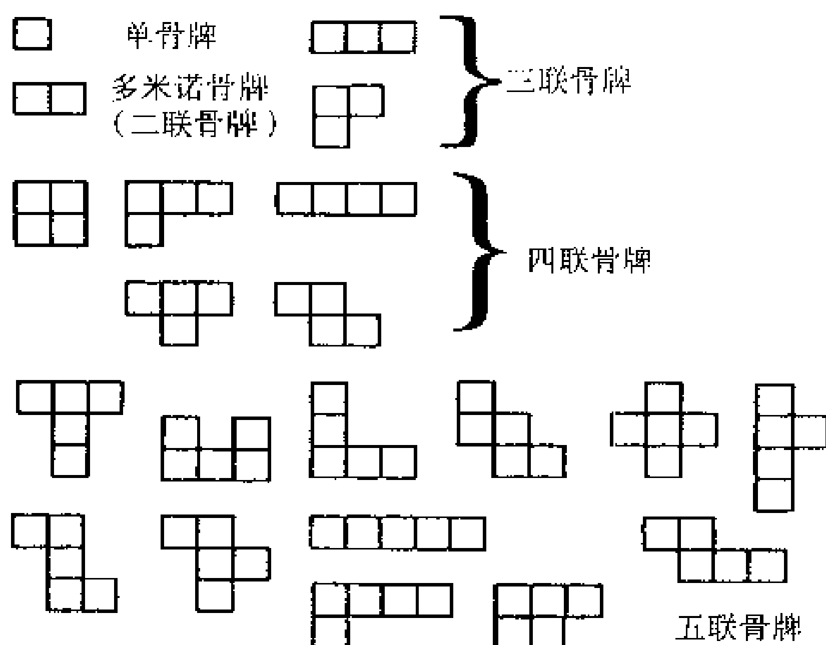


图 21.3 较简单的多联骨牌

① "多米诺"是domino的音译,其词头d本无独立意义. 但多米诺骨牌由两块正方形骨牌拼成,而英文中前缀di-表示"二"、"双"的意思,因此人们就把这个d看成是这样的意思,从而分离出"词干"omino,意思当然就是"正方形骨牌". 于是就有了polyomino(多联骨牌)、monomino(单骨牌)等名称. ——译注

多联骨牌理论中的一个典型问题是,确定哪些多联骨牌具有这样的性质:无限量地使用其一块特定多联骨牌的副本能够铺砌整个平面,或铺砌平面的一个象限,或铺砌以两条平行直线为边界的无限带形区域等等.(图 21.2 显示了是怎样用一块特[172]定的多联骨牌来铺砌一个带形区域,从而铺砌平面的.)可以设置一些限制条件,在铺砌中只可以根据所设的限制条件对基本图形使用旋转和/或镜射.可以研究这样的铺砌方式,其中允许用两种、三种或者其他数目的不同形状的铺砖.人们可以探寻这种形状的铺砖:用其几个相同的副本可以拼成原来形状的[173]一个放大的比例模型,或称**仿样**(replica),如图 21.4 所示.(许多年以前,即在 1962 年,我为这种形状的铺砖杜撰了一个专门名词:**仿样铺砖**(rep-tile).)

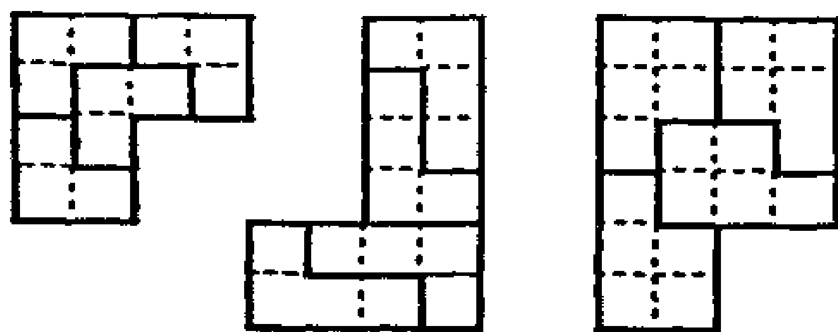


图 21.4 三种"仿样铺砖":3 联骨牌、4 联骨牌和 5 联骨牌各一

所有这些问题都已经得到研究.不过,本章首先着重讨论这样一个问题:在允许任意使用旋转和镜射的情况下,哪种形状的多联骨牌具有可以把其基本图形的某个有限数目的副本拼起来形成一个矩形的性质.这个问题是富有挑战性的,因为任意给出一块 n 联骨牌,对于形成一个矩形所**可能**需要的副本的最小数目,不可能设置一个推定的界限.也没有一个增长得足够快的关于 n 的精确表达式,例如 n^{n^n} 那样的,可以保证当基本图形的这么多副本的所有可能的拼搭都检查完毕而没有找到什么

矩形时，就不会存在包括比这还要多的副本的矩形．这些说法的根据就是前面提到的结果，即一**块任意的**多联骨牌是否能铺砌平面这个**一般性**问题是"计算不可判定的"．幸运的是，在通常遇到的**具体**情况中，答出这问题的可能性还是很大的(虽然没有十分把握)．但是不会有什么实验手册上的现成配方，也不会有什么操作手册上的既定程序，可以用来按部就班地指明一个给出的多联骨牌图形是否能铺砌某个(可能是巨大的)矩形．而正是因为像这样缺乏一个一般的判定程序，这个研究才如此令人感兴趣，如此富有挑战性．

多联骨牌的阶

1968年，克拉纳(David A. Klarner)把一块多联骨牌 P 的可拼成一个矩形(允许平移、旋转和镜射)的全等副本的最小数目定义为它的**阶**．对于那些不能铺砌任何矩形的多联骨牌，其阶没有定义．一块多联骨牌当且仅当它本身为一个矩形时阶为1．

一块多联骨牌当且仅当它是"一个矩形的一半"时阶为2，因为它的两个相同副本必须构成一个矩形．在实际操作上，这意味着这两个副本在构成一个矩形时互为180°旋转象．图21.5显示了一些例子．

不存在阶为3的多联骨牌．(这是由英格兰沃里克大学的斯图尔特(Ian Stewart)证明的．)事实上，能把任何一个矩形划分为一个"良态"几何图形的三个相同副本的唯一方法是把它分割为三个矩形(见图21.6)，而根据定义，一个矩形的阶为1．　[174]

能把四块相同的多联骨牌组合起来形成一个矩形的方法有

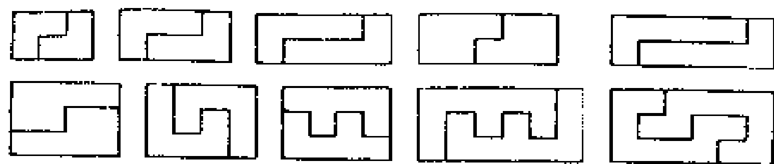


图21.5　一些2阶多联骨牌

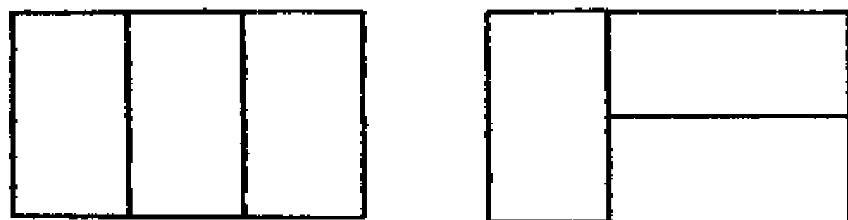


图 21.6　能把三个相同的矩形拼成一个矩形的方法

各种各样. 图 21.7 表明了一种方法,它是用一个图形的 4 个 90°旋转象来构成一个正方形.

另一种把 4 个相同图形组合起来形成矩形的方法利用了这矩形本身的四重对称性:上下对称、左右对称和 180° 旋转对称. 这方面的一些例子出现在图 21.8 中.

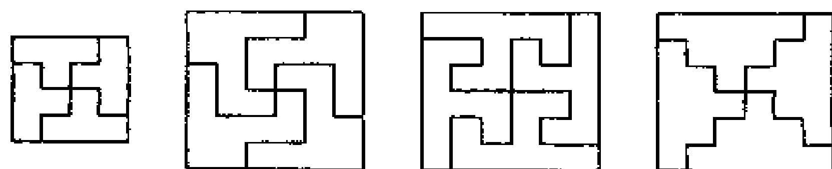


图 21.7　90° 旋转下的 4 阶多联骨牌

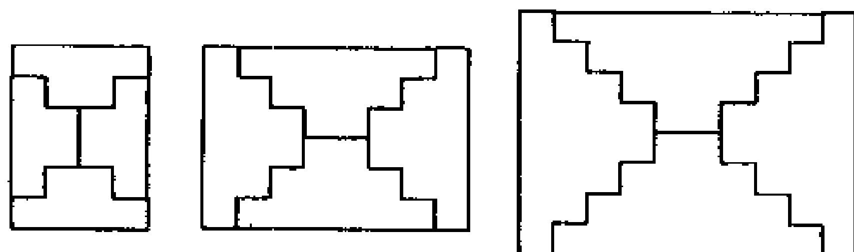


图 21.8　矩形对称下的 4 阶多联骨牌

还有更复杂的 4 阶样式,它们是克拉纳发现的,其中的两个在图 21.9 中例示.

关于 4 阶以上的情况,有着一种由戈隆布于 1985 年发现的系统构造方法,它对每一个正整数 s 都给出了阶为 $4s$ 的例子.

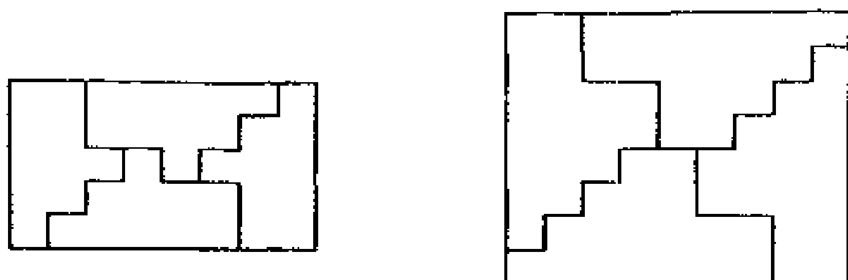


图 21.9　由克拉纳给出的其他4阶构造

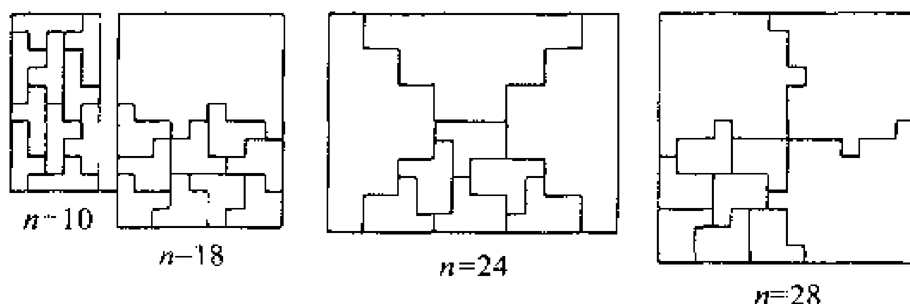


图 21.10　四个"零散的"多联骨牌,阶分别为10,18,24和28

人们已经知道了11个关于小多联骨牌的孤立例子,它们的阶分别为10,18,24,28,50,76,92,96,138,192和312.　[175]

图21.10显示了阶为10的孤立例子(戈隆布,1966)和阶为18,24及28的孤立例子(克拉纳,1969).

图21.11显示了阶为50的例子,它是由新西兰达尼丁的马歇尔(William Rex Marshall)于1990年发现的.

图21.12显示了阶为76和92的例子,它们都是由达尔克(Karl A.Dahlke)于1987年发现的.我在"关于趣味数学的施特伦斯[1]纪念会议"(卡尔加里,1986)上的讲话中提到过这两个问

[1] 施特伦斯(Eugène Strens),已故荷兰工程师,业余数学家,以表现数学而闻名的版画家及舍尔的朋友.生前大量收藏趣味数学书籍、剪报、杂志和手稿.——译注

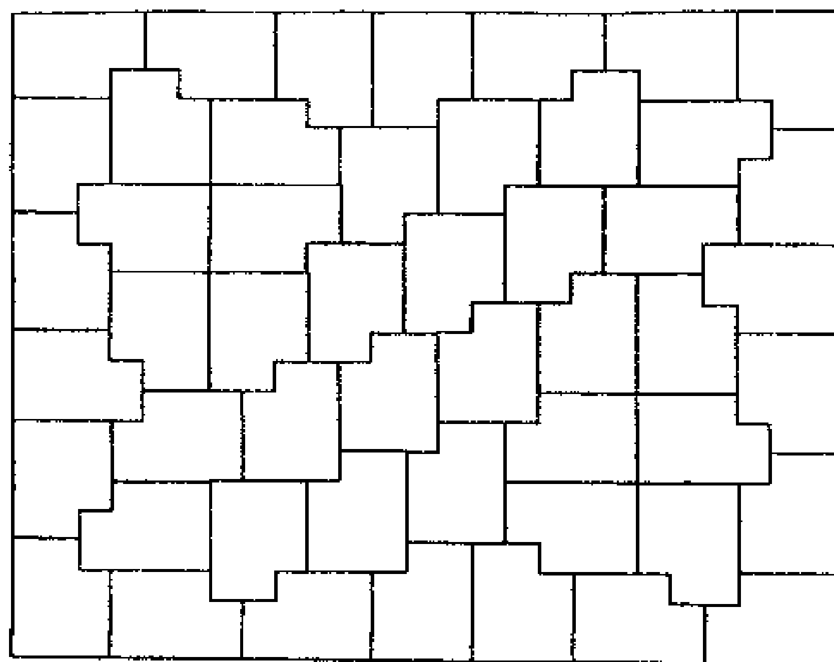


图 21.11　一种 50 阶的 11 联骨牌

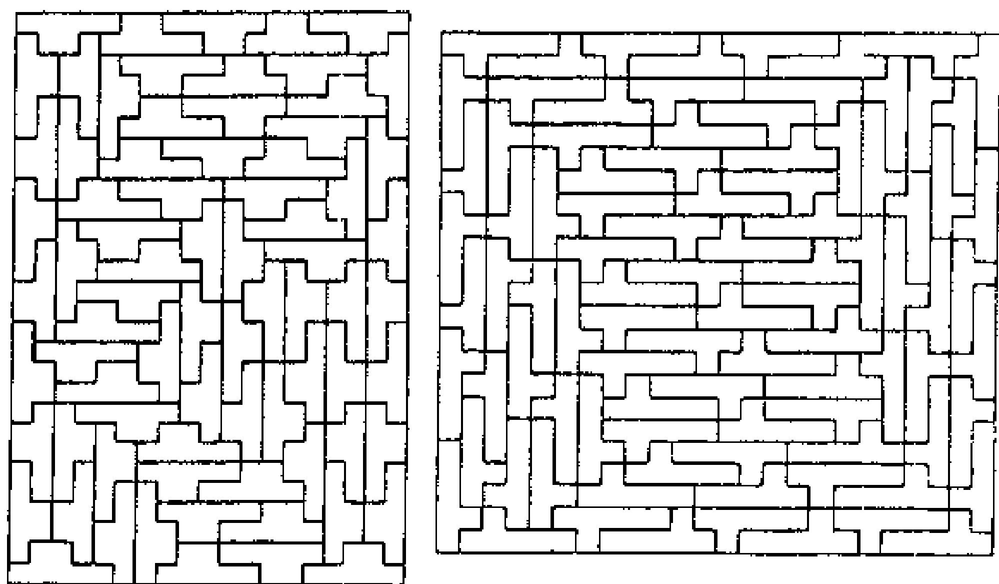


图 21.12　一种 76 阶的 7 联骨牌和一种 92 阶的 6 联骨牌

题，而彼得森[1]把它们写进了他在《科学新闻》上关于施特伦斯会议的报道之中．当达尔克给我寄来解答时，他说他找到它们只用了一台个人计算机．我通知了彼得森．彼得森便去采访达尔克，结果发现他竟是位盲人．我后来听说这两种铺砌实际上早在1985年就被马洛(T.W.Marlow)发现了．

图21.12中那种76阶7联骨牌在铺砌它的最小矩形时不能利用180°旋转对称性．图21.13中的96阶10联骨牌也是这样．它所铺砌的最小矩形(30 × 32)是由马歇尔于1991年发现的．1995年马歇尔还发现了192阶8联骨牌的最小矩形(32 × 48)和

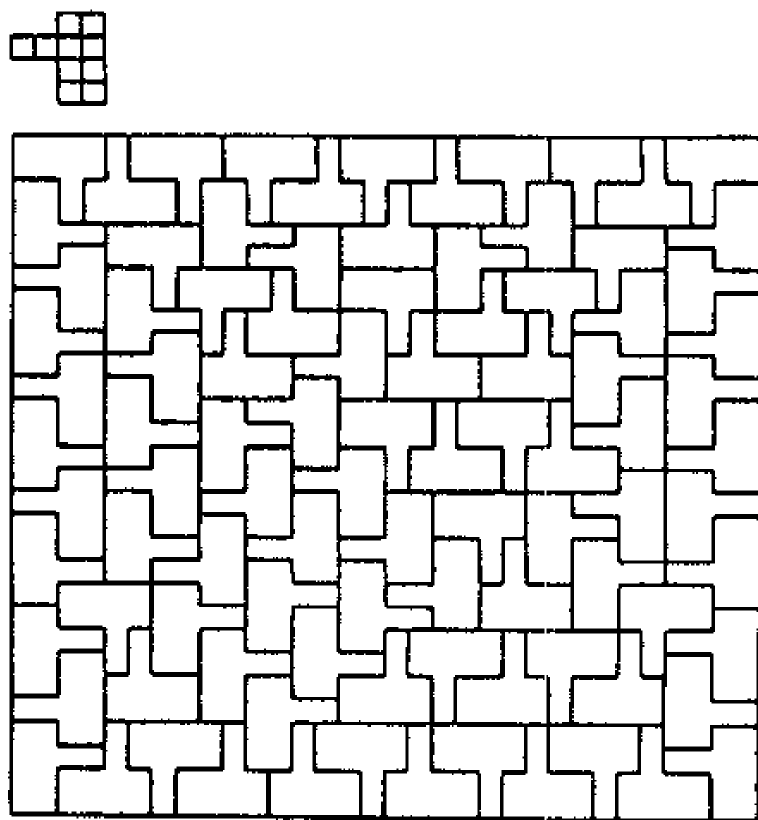


图21.13　一种96阶的10联骨牌

① 彼得森(Ivars Peterson)，美国《科学新闻》(*Science News*)杂志数学和物理方面的专栏作家．——译注

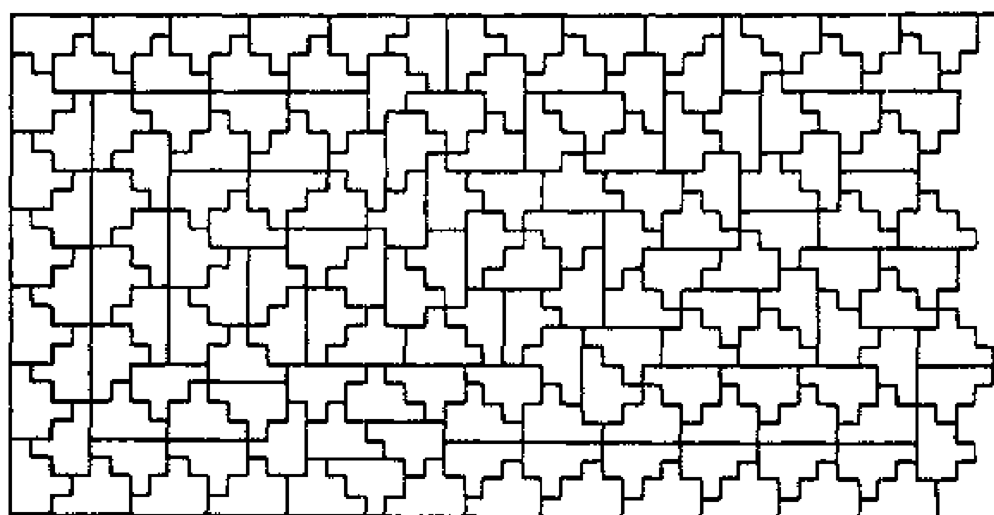


图 21.14　一个 312 阶的例子

[176] 138 阶 10 联骨牌的最小矩形(30 × 46). 它们将在后面考虑.

最后,图 21.14 显示了 312 阶的例子(达尔克,1988),虽然在这个情况中还不能绝对肯定不会存在由这种 8 联骨牌的更小数 [177] 目的副本所构成的矩形.

从未发现有阶为大于 1 的奇数的多联骨牌,但是存在这种多联骨牌(阶大于 3)的可能性也没有被排除.

已知的**偶数**阶多联骨牌除了阶为 2,10,18,50 和 138 外,其他的阶就是 4 的所有倍数. 奇怪的是,那些不是 4 的倍数的偶数阶都是 8 的倍数加 2. 是不是存在其他的偶数阶,它们可能是哪些数,这些仍不得而知. 还没有已知例子的最小偶数阶是 6 阶. 图 21.15 显示了一种用一块多联骨牌的 6 个副本配起来形成一个矩形的方法,但是这块多联骨牌(如图所示)实际上阶为 2. 里德(Michael Reid)最近找到了一种 6 阶的 7 **联响铃**①(一种用 7 个

① 原文为 heptabolo,从 diabolo 而来. 后者指一种玩具,类似我国的空竹,也称响铃. 与前一样,分离出 abolo,加上 hept-(七)等,便杜撰出一系列新词. ——译注

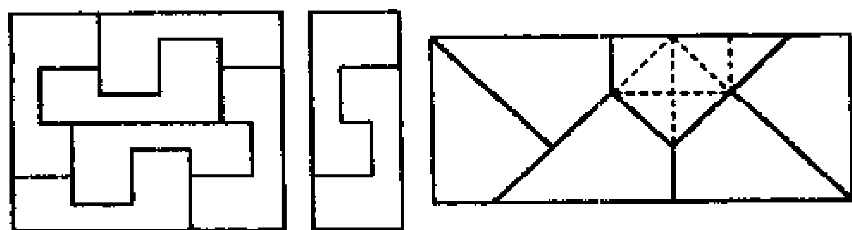


图 21.15 一种提示一种6阶铺砌的2阶12联骨牌，以及里德的6阶"7联响铃"（有6阶的多联骨牌吗？）

全等的直角等腰三角形组成的图形），也显示在图 21.15 中.

戈隆布的系统构造方法

戈隆布对 $4s$ 阶多联骨牌的构造方法当 $s=2$ 时给出了它第一个新例子，即一种 8 阶多联骨牌. 怎样把 8 个全等图形配合起来构成一个矩形的基本铺砌思想表示在图 21.16 中.

虽然图 21.16 中所用的图形不是多联骨牌，但这一思想可以用图 21.17 所示的 12 联骨牌来实现.（图 21.16 中所用的图形是一种 3 **联响铃**！）

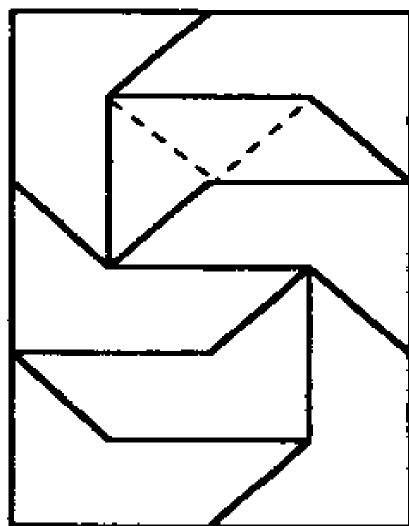


图 21.16 由 8 个全等块构成的一个矩形

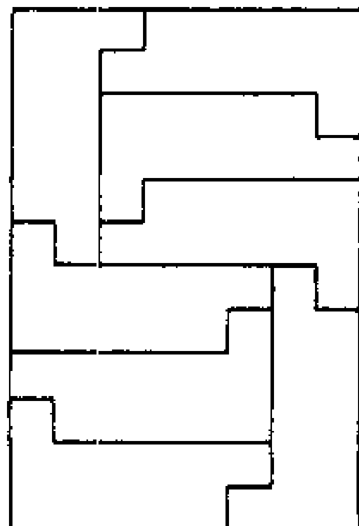


图 21.17 一种 8 阶的多联骨牌

要证明存在无穷多种互不相似的 8 阶多联骨牌，我们构造
178
了如图 21.18 所示的多联骨牌族．对每个整数 $r \geqslant 1$，这种构造方法都能产生一种 8 阶的 $(3r^2 + 6r + 3)$ 联骨牌，而且显然其中
179
任何两种都不相似．

同样容易证明，这些多联骨牌中没有一种能具有小于 8 的阶．这证明是这样的：首先注意到只有这"靴子的后跟"才能处于要铺砌的矩形的一个角，其次是这靴子的"鞋尖"必须与另一只靴子后上沿的缺口匹配．于是完成这个矩形的最快捷方法就需要这多联骨牌的 8 个副本．

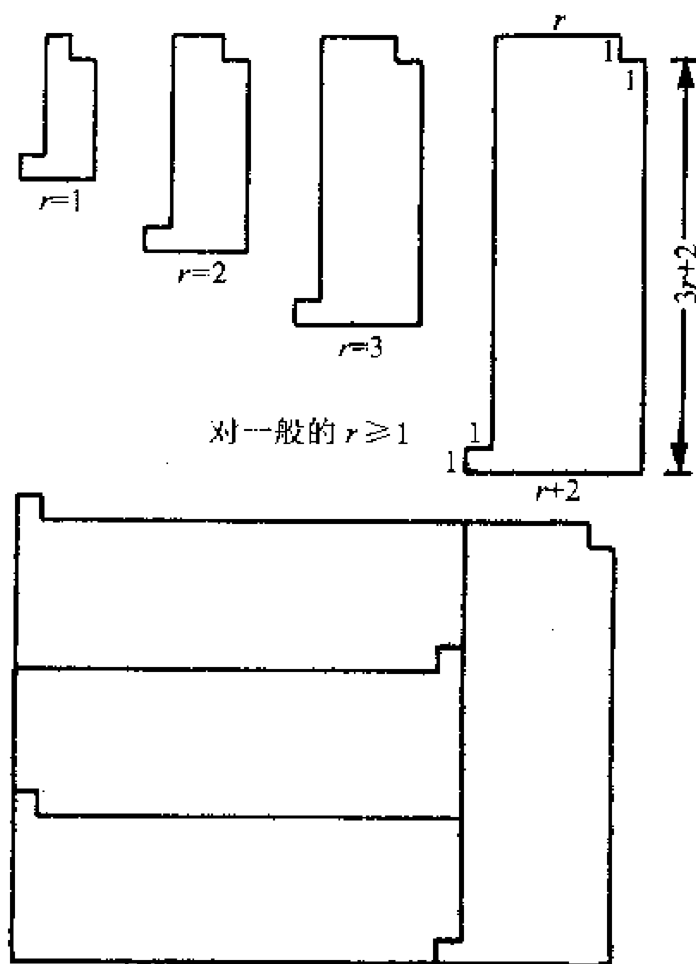


图 21.18　互不相似的 8 阶多联骨牌以及怎样堆放它们

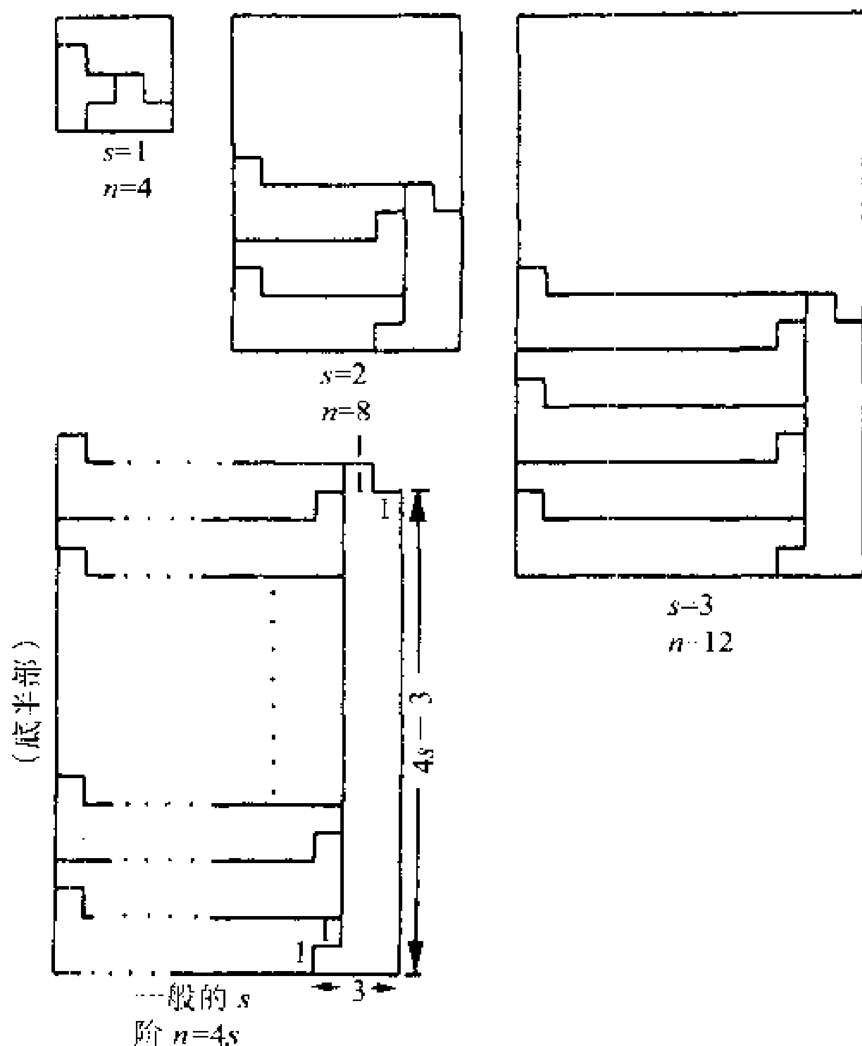


图 21.19 对每个正整数 s 都有一种阶 $n = 4s$ 的多联骨牌

在图21.19中，我们看到了一种对每个 $s = 1,2,3,4,\cdots$ 都能构造一种阶 $n = 4s$ 的多联骨牌的方法.(从一个 $2 \times (4s - 2)$ 的矩形开始，我们从一个角上取走一个正方形，再把它贴在斜对面的角上作为"鞋尖"，便得到 $4s$ 阶的多联骨牌.)

图21.18所表现的想法不仅可用于阶 $n = 8$ 的情况，还可用于任何阶 $n = 4s$ 的情况，从而得到无穷多种互不相似的 $4s$ 阶多联骨牌. 可把这种构造方法推广为涉及两个参数 r 和 s. 它从一

个$(r+1)\times(2s-1)(r+1)$的矩形开始,将一个1×1的正方形从这“靴子”的后上沿取走,成为其斜对角上的“鞋尖”.(关于这样得到的图形确实具有阶$n=4s$的证明类似于对$n=8$给出的证明.)

现在读者可能已经明白,多联骨牌学家们所玩的游戏是这样的:给出一块多联骨牌,它能铺砌吗?难道它不能铺砌?最近几年来,每当我公布一种其铺砌能力尚未得到确定的具体多联骨牌时,就有人带着一个不错的计算机程序前来,通常在一年之内,便有了一个铺砌矩形的解答.这次,我提出的是一个无穷的多联骨牌族,它的前几个已知能铺砌矩形,而且整个族中每四个就有一个能铺砌矩形.你能不能证明其他的**一种**或**所有**多联骨牌能铺砌矩形?这个多联骨牌族由图21.20例示.两种由马歇尔于1995年发现的最小矩形显示在图21.21中.

[180]

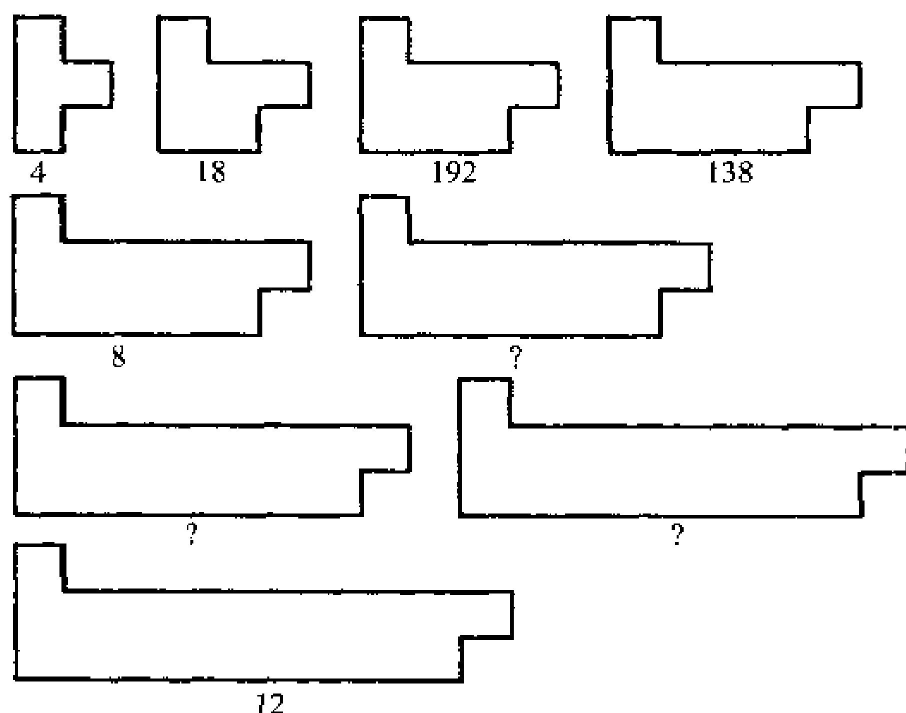


图21.20　多联骨牌的无穷族.每一种铺砖都能铺砌一个矩形吗?(每个图形下面的数字是相应的阶,如果它是已知的话.)

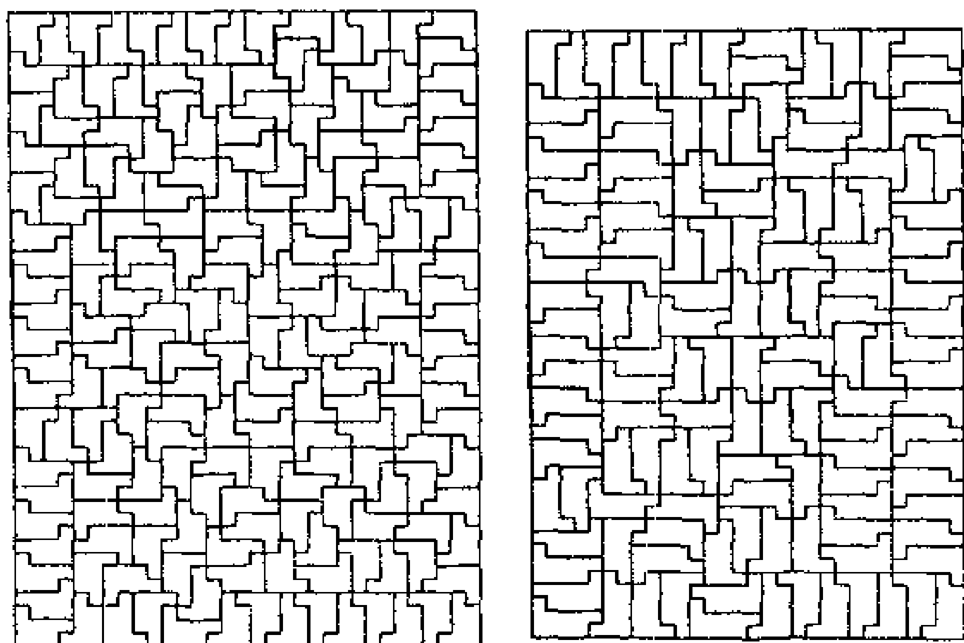


图 21.21 一种192阶的8联骨牌和一种138阶的10联骨牌

用多联骨牌铺砌其他形状的区域

现在我们回到除矩形之外的图形的铺砌问题．如果一块多联骨牌没有阶(也就是说如果它不能铺砌任何矩形)，它可能仍然能铺砌整个平面，或者铺砌平面的各种子区域，例如一个无限带形或一个弯折带形．这样的铺砌在图21.22中例示，用的分别是X形5联骨牌、F形5联骨牌和N形5联骨牌．

在图21.23中，我们看到了关于多联骨牌的一个铺砌分层结构(tiling hierarchy)(戈隆布，1966)．一种能铺砌这分层结构中所列出的某个区域的多联骨牌，也能铺砌其中所有较之为低的区域．于是，一种多联骨牌所属的"真正类型"(true catagory)就是这分层结构中它所能占据的最高方框．本章大部分内容是关于占据其中最高方框的那些多联骨牌的——也就是说，它们是那些能铺砌矩形的多联骨牌．已知具有成员的"真正类型"有"矩形"、"弯折带形"、"带形"、"本身"、"平面"和"什么也不能铺

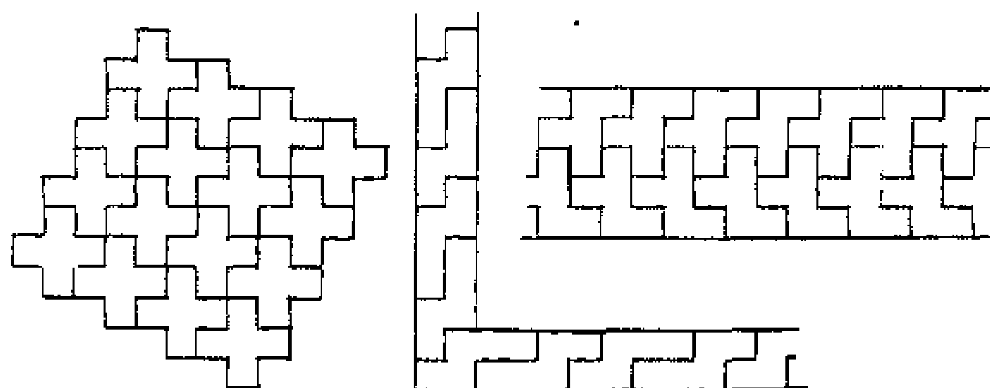


图 21.22 X 形 5 联骨牌铺砌了平面，F 形 5 联骨牌铺砌了一个带形，N 形 5 联骨牌铺砌了一个弯折带形

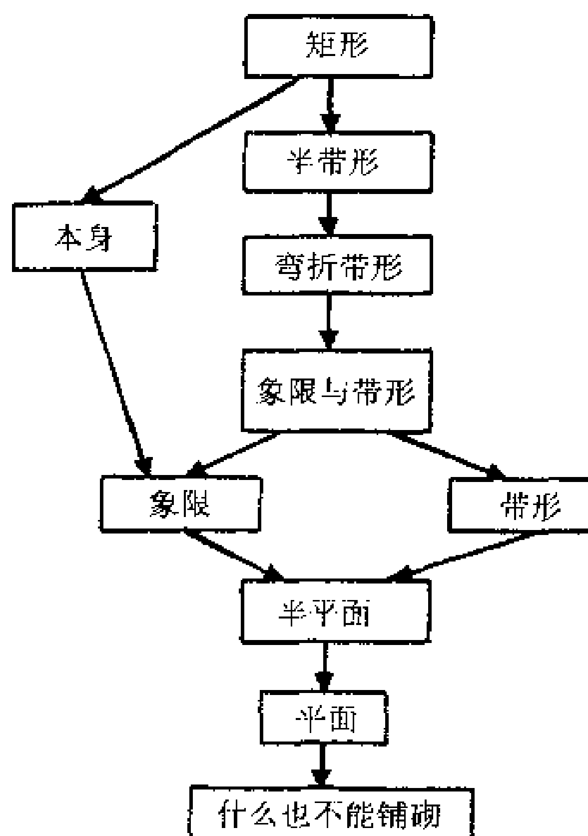


图 21.23 关于多联骨牌铺砌能力的分层结构

砌". 图 21.24 显示了类型"什么也不能铺砌"中的一个例子. (上述其他类型都已有例示. 图 21.4 中的仿样铺砖例示了类型"本身".) 对这分层结构中其他每个位置,是否有多联骨牌以这个位置作为自己所属的"真正类型",还是一个未解决的问题.

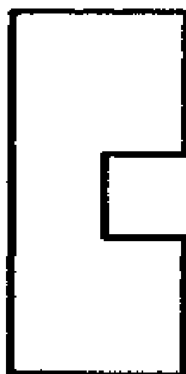


图 21.24 一种不能铺砌平面的多联骨牌

图 21.23 中的大多数包含关系是显而易见的. 要看出一种能铺砌一个弯折带形的多联骨牌既能铺砌一个象限也能铺一个平直带形,我们首先注意到弯折带形(如图 21.22 所示)可以被"嵌套"起来去覆盖平面的一个象限. 在进入最后一个重要话题后,我们将显示怎样从弯折带形走到平直带形.

关于平面铺砌,最具挑战性的未解决问题涉及这样一种图形,它能铺砌无限的平面,但只能用非周期的方式. 鲁滨逊率先找到了一个图形的集合(但是非常大),人们可以无限量地使用这集合中元素的副本来铺砌这无限的平面,但只能用非周期的方式. 他最终把这集合的大小缩减到只有 6 个元素. 彭罗斯(Roger Penrose)独立地发现了一个由6个图形组成的集合,它具有这种只能用非周期方式完成铺砌的性质. 他最终缩减到一个由两个图形组成的集合. 我相信希尔伯特那张名单的2000年版本会包括这样一个问题:是否存在一个单一的几何图形 S,使得人们可以用 S 的全等副本来铺砌无限的平面,但只能用非周期的方式? [182] [183]

为了更好地理解这个问题，请注意如果我们考虑的是一个无限的带形而不是平面，甚至我们允许使用多种形状的铺砖，那么回答就是否定的．要明白这一点，考虑一个水平的带形．然后在上边界上任意选取一个铺砖顶点，并按下述规则走出一条“曲折边界道路”：如果前面有一条向下的铺砖边界，就沿着它走；如果没有，就向右走．不断执行这种“向下第一，向右第二”的策略．有时候我们甚至可能被迫围着一块铺砖的突出部“向上”或“向左”行进，但是显然我们最终将抵达这带形的底边．于是我们可以很容易地根据这带形的厚度和所用各种铺砖的大小形状而得到一个关于所需“步”数的一致上界．这就意味着只可能存在有限多种形状不同的曲折道路，因此一定存在两个不同的起始顶点，它们给出了形状相同的道路．但是这样一来，如果我们以这两条相同道路之间的区域为基本图形取副本，把这些副本首尾相接，就得到了这个带形的所期望的周期性铺砌．而且我们看到，正如曾允诺要证明的，能铺砌一个弯折带形的多联骨牌也能铺砌一个平直带形（把上面的证明应用于那弯折带形的一条半无限臂即可），而且事实上这铺砌可以是周期性的．

在图 21.25 中，我们看到那“P形 5 联骨牌”是怎样用来非周期地铺砌一个象限的，方法是对这 P 形 5 联骨牌被分为其本身的较小仿样的仿样铺砖剖分进行迭代．我们把图形放大，重复这种剖分，再把图形放大，再重复这种剖分．“最终”，我们就把平面的这第一象限给填满了．利用关于坐标轴的镜射即可非周期地覆盖整个平面．（这并没有解决前面提到的问题，因为这种 [184] P 形 5 联骨牌也能周期性地铺砌平面.）

不仅这种 P 形 5 联骨牌，而且**每一种**多联骨牌仿样铺砖都能铺砌一个象限．为证明这点，令 J 是任意一块多联骨牌仿样铺砖，令 R 是包含 J 且其边平行于 J 中网格线的最小矩形[1]．我

[1] 注意这里的最小矩形与前面所说的不同．—— 译注

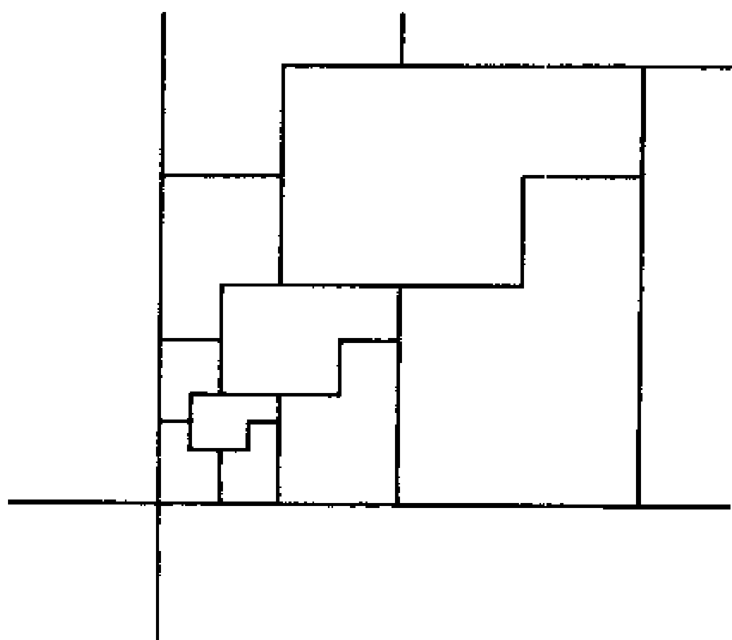


图 21.25 用P形5联骨牌的全等副本对一个象限的一种非周期仿样铺砖式铺砌

们首先证明一条引理：J 必定覆盖它这个最小矩形 R 的四个顶点中的至少一个. 假设不是这样. 在图 21.26 中，我们看到一块没有占据其最小矩形任何一个顶点的多联骨牌 K. 考虑 K 的在其最小矩形底边最左面的小正方形，它已用一个星号标出. 如果 K 是一块仿样铺砖，我们就可以把它分成全等的仿样，而且我们可以对这个过程进行迭代，直到每块仿样小得可以放进初始网格中的单单一个小方正方形. 到这一步，考虑 K 的覆盖了小正方形 * 左下角的仿样 k. 它完全位于那小正方形之中，并填满了小正方形的一个角，因此 k 显然占据了其最小矩形的一个角. 但是这同样的结论对 K 也是成立的.

现在，知道了 J 必定占据其最小矩形的至少一个角，就把 J 的这样一个角上的正方形放入平面第一象限的左下角. 然后对 J 进行仿样铺砖剖分，再放大到原来的尺寸，不断进行这种剖分

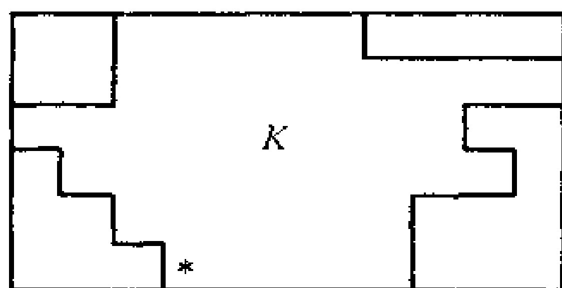


图 21.26　内接于其最小矩形的一块多联骨牌

和放大(就像我们对图 21.25 中的 P 形 5 联骨牌所做的那样),直到我们把整个第一象限都填满.(我们并不需要动用柯尼希(König)引理[①]来断言一个极限铺砌的存在.我们这个方法是构造性的.给出任意一个半径 R,不管它有多大,我们总可以把这种对第一象限的铺砌从原点出发明确地规定出来,直到距离为 R 的地方,而且这种铺砌样式不会随着 R 的增大而发生变化.如果离原点最近的 J 在一次迭代中位置被移动,这就可能要对这仿样铺砖剖分迭代过程采取每进行 m 次迭代仅作一次考察的方式[②].)

关于一种能铺砌一个弯折带形的多联骨牌也能铺砌一个平直带形和一种能铺砌本身的多联骨牌也能铺砌一个象限的证明,是由戈隆布给出的(1966).

---

① 柯尼希引理,又称柯尼希无穷性引理或无穷性引理.它有多种等价的表述,但译者欣赏这样一种比喻性的表述:假定人类不会灭绝,则必定存在一个人,他(她)的子孙后代能一直繁衍下去.它在数学上的严格表述及其对铺砌问题的应用可参见《数理逻辑通俗讲话》(王浩著,科学出版社,1981 年版).——译注

② 在图 21.25 中,每作一次迭代,原点处的仿样铺砖不但形状(当然)不变,而且方位不变.因此每次迭代后,原点附近的铺砌样式不变.但情况并不总是这样.有时经一次迭代后,原点处的铺砖在方位上发生了变化(例如把这块 5 联骨牌的另一个可与其最小矩形顶点相合的顶点放在原点处时),从而铺砌样式也发生了变化.这就同前一句话发生矛盾.但是一块仿样铺砖在原点处的不同方位只有有限多种,因此必定存在一个正整数 m,使得每经过 m 次迭代,原点处铺砖又恢复了原来的方位,从而恢复了原来的铺砌样式.——译注

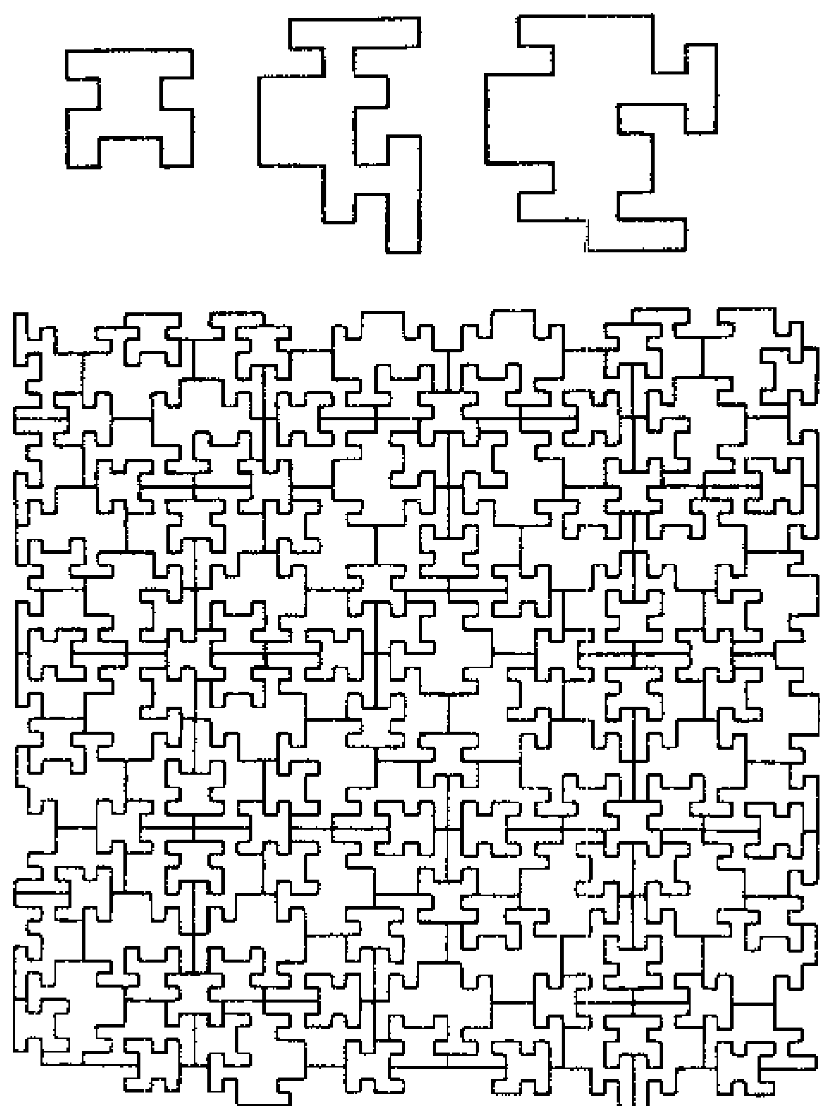


图 21.27　彭罗斯的由三块多联骨牌组成的集合，能用来铺砌平面，但不能作出周期性铺砌

最后，回到平面的非周期铺砌这个话题．就我们现在所介绍的内容，要问的问题自然是，是否存在用多联骨牌作出的"不可周期化的"铺砌．当前最好的结果是如图 21.27 所示的那个由三个图形组成的集合．这个集合能用来铺砌平面，但是不能

作出周期性铺砌.（这个集合是最近由彭罗斯发现的，他写道，这是从安曼(Robert Ammann)的一个铺砌集合得来的.）彭罗斯的异周期性铺砌还同"准晶体"(quasicrystal)有着联系，后者近年来引起化学家的很大兴趣.

某些"块"集合能铺砌平面，但不能作出周期性铺砌. 有关[185]的证明既机智又精妙(见[10]的第10章)，而且无疑超出了希尔伯特曾经仔细思考过的范围. 但是希尔伯特把一个关于铺砌和填装的问题包括进他那张著名的名单之中，可见他洞察力之深邃. 这是一个对业余爱好者来说可以涉足的课题，但是它离数学的心脏又是如此之近，而且不断地提供着似乎无穷无尽的问题，这些问题既引人入胜又令人振奋.

注记 这些材料基于戈隆布所著的《多联骨牌(修订版)》(*Polyomino—Revised Edition*, Princeton University Press, 1994)的第8章. 这是对《多联骨牌(第一版)》(*Polyomino*, Scribner's Sons, 1965)内容所增加的新的一章.

参考文献

1. Robert Berger, The undecidability of the domino problem, *Memoirs* [186] *of the American Math. Society* 66(1966), 1—72.
2. Karl A. Dahlke, The Y-hexomino has order 92, *Journal of Combinatorial Theory, Series A* 51(1989), 125—126.
3. Karl A. Dahlke, A heptomino of order 76, *Journal of Combinatorial Theory, Series A* 51(1989), 127—128.
4. Karl A. Dahlke, Solomon W. Golomb, and Herbert Taylor, An octomino of high order, *Journal of Combinatorial Theory, Series A* 70 (1995), 157—178.
5. Martin Gardner, Mathematical Games: On "rep-tiles", polygons that can make larger and smaller copies of themselves, *Scientific*

American 208(1963),154—164.

6. Solomon W. Golomb, Replicating figures in the plane, *Mathematical Gazette* 48(1964),403—412.
7. Solomon W. Golomb, Tiling with polyominoes, *Journal of Combinatorial Theory* 1(1966),280—296.
8. Solomon W. Golomb, Tiling with sets of polyominoes, *Journal of Combinatorial Theory* 9(1970),60—71.
9. Solomon W. Golomb, Polyominoes which tile rectangles, *Journal of Combinatorial Theory, Series A* 51(1989),117—124.
10. B. Grünbaum and C. C. Shephard, *Tilings and Patterns*, New York: Freeman, 1987.
11. A. S. Kahr, E. F. Moore, and H. Wang, Entscheidungsproblem reduced to the ∀ ∃∀ case, *Proceedings National Academy of Science USA* 48(1962),365—377.
12. David A. Klarner, Packing a rectangle with congruent N-ominoes, *Journal of Combinatorial Theory* 7(1969),107—115.
13. William Rax Marshall, private communications dated 14 May, 1990, 25 November, 1991, and 6 June, 1995.
14. Roger Penrose, *Shadows of the Mind*, Oxford University Press, 1994.
15. Michael Reid, private communications from 1992 to 1995.
16. Ian Stewart and A. Wormstein, Polyominoes of order 3 do not exist, *Journal of Combinatorial Theory, Series A* 61(1992),130—136. [187]

[188]

第22章 一个模式问题,一个概率论悖论和一个美妙的证明

问题与模式

设 S 是一个非负整数的有限集合,我们定义 mex{S} 为不是 S 中元素的最小非负整数.(mex 是 minimum-exclude(最小排除)的缩写(见伯利坎普、康韦和盖伊的《取胜之道》).)

下面所说的二维阵列 $\boldsymbol{A}$,假设可以无限地延伸出去,并满足下述这个条件:

(*)第 i 行第 j 列的元素 a_{ij} = mex{S_{ij}},其中 S_{ij} 是由第 i 行中位于该元素左面的和第 j 列中位于该元素上面的元素所组成的集合.

问题 *第 777 行第 1001 列的元素是什么?*

(当然,我们实际上是要求给出一个关于第 i 行第 j 列元素的精确表达式.)

显然有一个且只有一个阵列满足(*),具体地说,a_{00} 必定为零,从而元素 a_{01} 和 a_{10} 为1,等等. 而且我们看到,在构建这个矩阵的过程中,那一行行(或以其他任何方式)出现的元素,每一个都被"强迫就位". 只要你愿意,你可以不用求助于计算机技术而机械地算出一项又一项. 这里以及下面其他地方,我们把行和列的指标定为从0开始. 后来出现的情况令人十分意外. [189] 确实难以想像,如果不进行这种预备性的试验,人们怎么能猜出

有关的结果. 表 22.1 是这个阵列的一部分.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	0	3	2	5	4	7	6	9	8	11	10	13	12	15	14	17	16	19
2	3	0	1	6	7	4	5	10	11	8	9	14	15	12	13	18		
3	2	1	0	7	6	5	4	11	10	9	8	15	14	13	12			
4	5	6	7	0	1	2	3	12	13	14	15	8	9	10	11			
5	4	7	6	1	0	3	2	13	12	15	14	9	8	11	10			
6	7	5	4	2	3	0	1	14	15	12	13	10	11	8	9			
7	6	4	5	3	2	1	0	15	14	13	12	11	10	9	8			
8	9	10	11	12	13	14	15	0	1	2	3	4	5	6	7			
9	8	11	10	13	12	15	14	1	0	3	2	5	4	7	6			
10	11	8	9	14	15	12	13	2	3	0	1	6	7	4	5			
11	10	9	8	15	14	13	12	3	2	1	0	7	6	5	4			
12	13	14	15	8	9	10	11	4	5	6	7	0	1	2	3			
13	12	15	14	9	8	11	10	5	4	7	6	1	0	3	2			
14	15	12	13	10	11	8	9	6	7	4	5	2	3	0	1			
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0			
16	17	18																
17	18																	

表 22.1

这里正在发生什么? 表22.2对 $\boldsymbol{A}$ 所作的划分显示了正在发生的事情.

我们曾指出, a_{00} 必定为 0, a_{01} 和 a_{10} 为 1, 因而 a_{11} 又是 0. 如果我们把这 2×2 矩阵称作 A_{11}, 那么我们看到, A_{11} 右面和下面的邻接矩阵 A_{12} 和 A_{21} 就是在 A_{11} 的每个元素上加上 2 而得来的, 而 A_{22} 则是 A_{11} 沿对角线平移的结果. 我们现在有了一个 4×4 矩阵 A_{44}. 要得到 A_{44} 下面和右面的 4×4 矩阵, 我们把 4 加到 A_{44} 的元素上. 然后把 A_{44} 在对角线上复制出来, 这就给出了一个 8×8 阵列. 继续用这种方式, 就得到一个 16×16 阵列, 依此类推. 至此, 它的一般构造已经变得很清晰了. [190]

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	0	3	2	5	4	7	6	9	8	11	10	13	12	15	14	17	16	19
2	3	0	1	6	7	4	5	10	11	8	9	14	15	12	13	18		
3	2	1	0	7	6	5	4	11	10	9	8	15	14	13	12			
4	5	6	7	0	1	2	3	12	13	14	15	8	9	10	11			
5	4	7	6	1	0	3	2	13	12	15	14	9	8	11	10			
6	7	4	5	2	3	1	0	14	15	12	13	10	11	8	9			
7	6	5	4	3	2	1	0	15	14	13	12	11	10	9	8			
8	9	10	11	12	13	14	15	0	1	2	3	4	5	6	7			
9	8	11	10	13	12	15	14	1	0	3	2	5	4	7	6			
10	11	8	9	14	15	12	13	2	3	0	1	6	7	4	5			
11	10	9	8	15	14	13	12	3	2	1	0	7	6	5	4			
12	13	14	15	8	9	10	11	4	5	6	7	0	1	2	3			
13	12	15	14	9	8	11	10	5	4	7	6	1	0	3	2			
14	15	12	13	10	11	8	9	6	7	4	5	2	3	0	1			
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0			
16	17	18																
17	18																	

表 22.2

这个阵列那不同凡响的特征就是,2 的幂在这里起了完全没有预料到的作用. 条件(∗)中根本没有什么东西可以引导人们想到这一点. 不过,既然已经把它指出来了,那么我们自然要把整数用二进制记号写出,到这一步,秘密就被揭开了. 表 22.3 是这个阵列的开头部分.

0	1	10	11	100	101	110	111	1000
1	0	11	10	101	100	111	110	1001
10	11	0	1	110	111	100	101	1010
11	10	1	0	111	110	101	100	1011
100	101	110	111	0	1	10	…	

表 22.3

定义 设 i 和 j 都是整数,我们把它们的二进制展开看作 $\mathbf{Z}_2$ 上的向量. 如果一个整数 k 的二进制展开是 i 与 j 的向量和,那么我们就称 k 为 i 与 j 的**nim 和**,记为 $k = i \oplus j$. 这样,当且仅当 i 或 j 的第 r 位数字为1但不都为1时,k 的第 r 位数字为1.

我们采用 $\boldsymbol{A}$ 的行和列从0开始而不是从1开始标记的约定. 于是这故事如下: [191]

定理 $\boldsymbol{A}$ 中的 a_{ij} 就是 $i \oplus j$.

这个定理的一个有趣的推论是,$\boldsymbol{A}$ 是非负整数上一个阿贝尔群的乘法表,这是因为nim和把一个整数的 n 位二进制展开作为 $\mathbf{Z}_2$ 上 n 维向量空间的一个元素来处理. 人们可以验证 $\boldsymbol{A}$ 的元素满足结合律.

表22.1 ~ 表22.3为我们定理的正确性提供了非常令人信服的"证据",但是下面这个证明恐怕也不是完全平凡的.

证明 令 $i \oplus j = n$. 我们必须证明,若 m 小于 n,则或者存在 $i' < i$ 使得 $i' \oplus j = m$,或者存在 $j' < j$ 使得 $i \oplus j' = m$.

如果 d 是一个被看作向量的二进制数,则令 d_k 为它的第 k 位数字(从右数起). 现在令 r 是使得 $m_r \neq n_r$ 的最大下标. 那么 $n_r = 1, m_r = 0$,这是因为 $m < n$. 因此不是 i_r 为1就是 j_r 为1,不妨设 i_r 为1. 令 i' 是使得 $i' \oplus j = m$ 的(唯一)二进制数向量. 那么对 $k > r$,有 $i'_k = i_k$,而 $i'_r = 0$. 所以 i' 小于 i;于是 $m = a_{i'j}$ 在 a_{ij} 的左面.

最初那个要求找出第777行第1001列元素的问题答案如何?好吧,我们有

$$777 \to 1\ 1\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 1 \quad 和$$

$$1001 \to 1\ 1\ 1\ 1\ 1\ 0\ 1\ 0\ 0\ 1; \quad 所以$$

$$777 \oplus 1001 = 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0 = 224.$$

评注 对 n 维阵列的情况，这个定理完好无损；具体地说，(*)成为这样的条件：在位置 $(i_1,\cdots,i_k,\cdots,i_n)$ 上的元素是

mex{对所有的 k，在位置 $(i_1,\cdots,i'_k,\cdots,i_n)$ 上的元素，其中 $i'_k < i_k$}.

那么在位置 $(i_1,\cdots,i_k,\cdots,i_n)$ 上的元素结果是 $i_1 \oplus i_2 \oplus \cdots \oplus i_n$.

为什么那个术语要叫"nim 和"？好，考虑一个有着 n 堆棋子的 Nim 游戏，其中第 k 堆有 i_k 个棋子. 熟悉 Nim 游戏的人会知道，当且仅当 i_k 的 nim 和为 0 时，这个**局面**对第二位局中人来说是必胜的. 一般地，我们可以把 $i_1 \oplus i_2 \oplus \cdots \oplus i_n$ 看作这个 Nim 游戏的**值**，而我们的定理断言，一位局中人仅从一堆棋子中取子就能走到任何一个更小的值. 两堆的 Nim 当然是没有什么意思的，因为 0 全部出现在整条对角线上，但是 3 堆的情况就不同了，例如有 $1 \oplus 2 \oplus 3 = 0$.

人们可以通过改变集合 S_{ij} 的定义来对那个最初提出的问题创造一些变化形式. 一个有趣的变化形式是把对角线上向后的部分包括进 S，也就是说，扩充 S_{ij}，把所有满足 $i' < i, j' < j$ 且 $i' - j' = i - j$ 的 $a_{i'j'}$ 也包括进来. 对这种情况，关于 a_{ij} 的精确表达式尚不得而知，但是有一种方法可以把所有的 0 找出来. 这对应着一种 2 堆的类 Nim 游戏，其中每一步或者像普通 Nim 那样在其中一堆棋子中取走任意数量的棋子，或者在两堆棋子中取走相同数量的棋子①.

有一种 Nim 的变化形式尚未得到解决. 它就是普通的 Nim 再加上游戏过程中允许任一位局中人采用一次"放弃"策略，但有一位用了之后，另一位就不能用了，而且在所有的棋子都取完后即使没有人用过也不能用. 同样，最后取子的局中人为赢家.

① 此即威索夫(W. A. Wythoff)游戏，见本译丛中《数学游戏与欣赏》的第 1 章. ——译注

又一个概率论悖论

设有一个分布已知的随机变量 X,它从一有限序集 S 中取整数值. 例如,S 可以是从 1 到 10 的整数. 用 p_i 记 X 等于 i 的概率,用 D_i 记 X 大于或等于 i 的概率.

"游戏"(以大自然为对手)Γ^2 就是取 X 的独立观察值构成一个序列 $x_1,x_2,\cdots$,其中一旦有某个 x_i 成为"第二大",即在 x_i 前面只有一项大于或等于 x_i,这个序列立即终止. 这位观察者因此而获得一笔价值相当于这个 x_i 的**支付**.

游戏 Γ^k 与上面相同,只是"第二大"被"第 k 大"所代替.

游戏 Γ_k 与 Γ^k 相同,只是"第 k 大"被"第 k 小"所代替.

例子 对于序列$(3,7,2,5,1,6,1,5^*,9,2,\cdots)$,游戏 Γ^4 的支付是 5,这是因为前面恰有三个等于或大于 5 的观察值 x_2,x_4 和 x_6.

注记 在游戏 Γ_1 和 Γ^1 中,支付就是 x_1.

问题 在 $\Gamma^2,\Gamma^{17},\Gamma_2$ 等等游戏之中,哪一个对观察者来说最有利?

回答 它们全都一样.

定理 *Γ^k 和 Γ_k 的支付的概率分布与对 S 的原始分布相同.*

这种说法一看上去似与直觉相悖. 人们总觉得在"第二大"上打赌应该比在"第十七大"或"第二小"上打赌更有利些.

这个定理实质上是概率论中一个名叫伊格纳托夫(Ignatov)定理的结果的离散情况. 这个结果可通过直接计算来证明. 方法是在所有观察值序列的集合中考察可能的组合,并在出现的二项式系数中注意发现一些恒等式. 与此相反,我提出一个简短的间接证明,它基本上不需要计算. 下面这个非正式的证明

可以很容易地被正式化. 它依赖于 Γ^k 的支付的一个奇妙特性.

记号 给出一个观察值序列 $Z=(x_1,x_2,\cdots)$,令 Z_r 为 Z 中所有大于或等于 r 的项所组成的子序列.

术语 我们将把 Γ^k 的支付称为 **k 支付**.

引理 *序列 $Z=(x_1,x_2,\cdots)$ 的 k 支付就是 S 中使得 Z_r 的第 k 项为 r 的最小数 r*[①].

证明 如果游戏以支付 r 结束,这就意味着前面恰有 $k-1$ 个观察值大于或等于 r,但这就等于说 r 是 Z_r 的第 k 项. 最后,r 必定是具有这种性质的最小数,因为如果有一个更小的这样的数,这游戏在早些时候就结束了.■

[193]

我们的定理可以直接从这引理推出. 具体地说,Z_s 的第 k 项为 s 的概率是 p_s/D_s,不为 s 的概率是 D_{s+1}/D_s. 而且,如果 Z_s 的第 k 项不为 s, 那么它就大于 s, 但是这并没有给出关于 Z_{s+1} 的第 k 项的值的任何信息. 这样一来, 支付 r 的概率就是 $D_1(D_2/D_1)\cdots(D_r/D_{r-1})(p_r/D_r)=p_r$.

这定理对游戏 Γ_k 也成立这一事实可由对称性推出.

不过人们可以说得比这再多一些. 给出一个序列 Z,假定人们知道它的 2 支付 —— 比方说 —— 是 5,那么据此能推断出关于 3 支付的什么情况吗? 能推断出关于 17 支付的什么情况吗? 回答是:什么也不能推断出来. 换句话说,给出序列 Z,如果 j 不等于 k,那么 j 支付与 k 支付就是**独立的**. 这一点仍由引理得出,这引理表明 j 支付(k 支付)只取决于 $Z_1,Z_2,\cdots$ 这些序列的第 j 项(第 k 项);但是这些项是独立的,因为原来的观察是独立的. 更一般地说,知道了任何有限的值集合 F 中的支付,并不能给出关于不在 F 中的值的任何信息,这就是独立的意思.

① 这句话似有点费解,但看来也没有其他较简短的表述方法. 或许用一个式子反而明确:$r=\min\{s\in S:Z_s$ 的第 k 项 $=s\}$. —— 译者注

一个美妙的证明

在第 19 章中，霍尔顿为用最经济的方式系鞋带这个自然的"日常生活问题"找到了答案．下面这篇由米休列维奇(Michal Misiurewicz)撰写的说明性文章，给出了霍尔顿结果的一个变化和扩充．

为不规则的鞋子系鞋带

米哈尔·米休列维奇(Michał Misiurewicz)

霍尔顿考虑了用最有效的方式为一只鞋子系鞋带(以鞋带的长度为衡量标准)的问题．让我们回忆一下这个问题．我们有平面上的 $2(n+1)$ 个点(鞋眼)$A_0,A_1,\cdots,A_n,B_0,B_1,\cdots,B_n$．一种**系带法**是指一条道路 $C_0\to C_1\to\cdots\to C_{2n+1}$，其中 $\{C_0,C_1,\cdots,C_{2n+1}\}=\{A_0,A_1,\cdots,A_n,B_0,B_1,\cdots,B_n\}$，$C_0=A_0$，$C_{2n+1}=B_0$，而且集合 $\{A_0,A_1,\cdots,A_n\}$ 中的点与集合 $\{B_0,B_1,\cdots,B_n\}$ 中的点在这条道路上交替出现．它的**长度**是 $|C_0C_1|+|C_1C_2|+\cdots+|C_{2n}C_{2n+1}|$．这是鞋带被用到的长度，不包括松在外面的鞋带头子，虽然这是系紧鞋带所必需的．这个问题是要找出长度最短的系带法．

霍尔顿证明了**标准系带法**①就是这最短的系带法．这种标准系带法是：如果 n 为奇数，则按 $A_0\to B_1\to A_2\to B_3\to\cdots\to A_{n-1}\to B_n\to A_n\to B_{n-1}\to\cdots\to A_3\to B_2\to A_1\to B_0$ 系鞋带；如果 n 为偶数，则按 $A_0\to B_1\to A_2\to B_3\to\cdots\to B_{n-1}\to A_n\to B_n\to A_{n-1}\to\cdots\to A_3\to B_2\to A_1\to B_0$ 系鞋带(见图 22.1)．

本节的目的是对一个广泛得多的结果给出一个简短的证明．霍尔顿假设这些鞋眼排成平行的两行，相邻鞋眼之间也是

① 霍尔顿称这种系带法为**美国式**．然而，我想不起有什么人，不管在美国还是在欧洲，他或她的系鞋带方式会有不同．—— 原注

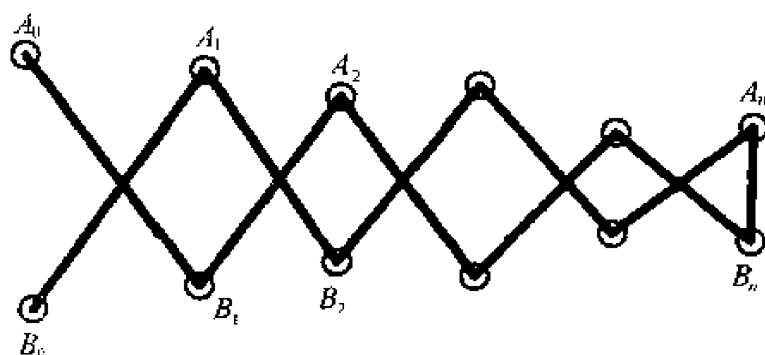
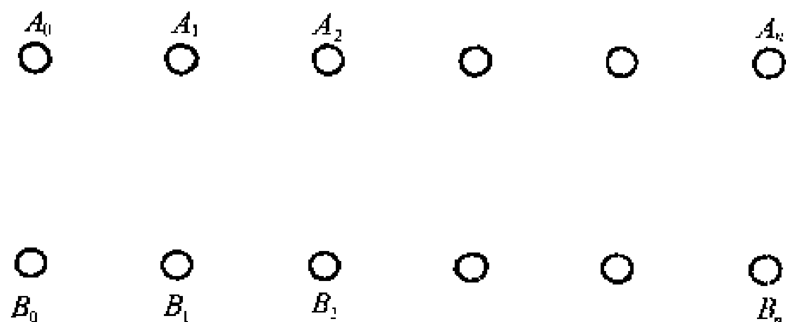


图 22.1 标准系带法

等距的. 更准确地说, 存在常数 $v, w > 0$, 使得在等高的水平上, 有 $A_i = (iv, w)$, $B_i = (iv, 0)$ (见图 22.2). 这是对现实情况的一个粗糙的近似(见图 22.3). 在现实情况中, 鞋眼之间的"水平"距离和"垂直"距离都可能有所不同. 甚至关于水平线的对称性也可能不存在, 特别是当鞋子被穿旧了的时候. [194]



[195] 图 22.2 鞋眼的霍尔顿式排列

我要在一个看来更贴近现实的假设下给出标准系带法最小性的一个证明. 这假设是:

(1) 对任何 k, l, 通过 A_k 和 B_l 的直线都把这鞋眼集合分为 $\{A_i : i < k\} \cup \{B_j : j < l\}$ 和 $\{A_i : i > k\} \cup \{B_j : j > l\}$.

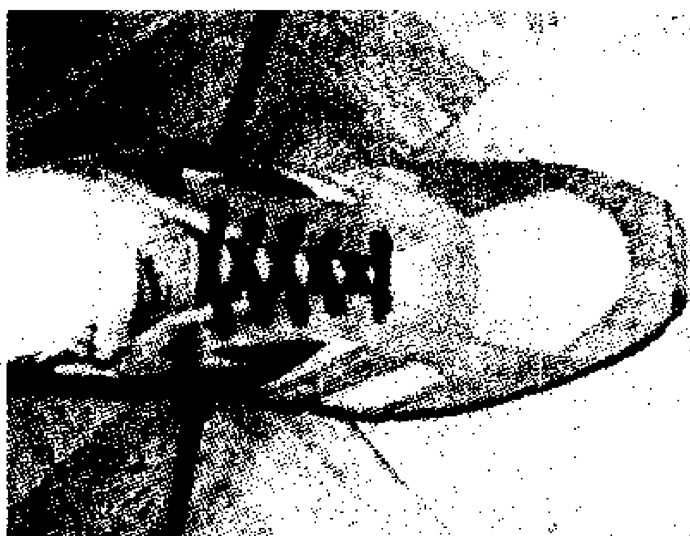


图 22.3　一只现实的鞋子

换句话说,如果我们从任何一个"B"鞋眼去观察"A"鞋眼(或从任何一个"A"鞋眼去观察"B"鞋眼),它们都以自然顺序排列:$A_0,A_1,\cdots,A_n$(或 $B_0,B_1,\cdots,B_n$).

事实上,甚至这个假设也太强了.我们只需要下面的

(2) 对任何 $i<j$ 和 $k<l$,有

$$|A_iB_k|+|A_jB_l|<|A_iB_l|+|A_jB_k|.$$

要看出从(1)能推出(2),取 $i<j$ 和 $k<l$,并假定(1)成立.于是,如图22.4所示,鞋眼 A_i 和 B_l 位于通过 A_j 和 B_k 的直线的两侧,而鞋眼 A_j 和 B_k 位于通过 A_i 和 B_l 的直线的两侧.因此,线段 A_iB_l 与线段 A_jB_k 相交.设 D 是交点,那么根据三角不等式,我们

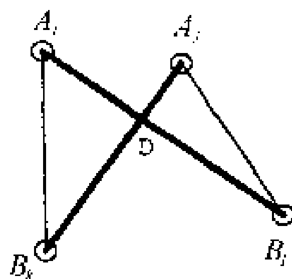


图 22.4　从(1)推出(2)

得到

$$|A_iB_k| + |A_jB_l| < |A_iD| + |DB_k| + |A_jD| + |DB_l| < |A_iB_l| + |A_jB_k|;$$

所以(2)成立.

我们可以考虑一种甚至更现实的模型,其中鞋眼是三维空间中的点.不过,鞋带应该系得紧贴鞋面.因此,我们可以假设鞋眼是鞋子曲面上的点,而距离应该沿着这个曲面测量.读者可以进行精确测量,以验证在这个模型中他或她的鞋子是否满足假设(2).但是通常鞋子曲面上与系鞋带有关的那一片曲率非常小.这意味着如果我们把这曲面看作是一片"被弄弯了的"平面,我们不会造成很大的误差.把它摊平,就使我们回到了我们原来的模型.现在只要迅速一瞥,就可以验证假设(1)是否被
[196] 满足.

到了证明本节主要结果的时候了.

定理 *假设鞋眼集合$\{A_0, A_1, \cdots, A_n, B_0, B_1, \cdots, B_n\}$满足(2),那么标准系带法比其他任何系带法都要短.*

这个证明的思想相当简单.仍请参见图22.4,让我们定义一次"移动"就是把一对线段A_iB_l和A_jB_k用A_iB_k和A_jB_l代替.唯一的问题是经这样一次移动后得到的结果可能不成为一种系带法.例如,如果我们从图22.5所示的系带法开始,进行一次把线段A_0B_2和$A_{n-1}B_1$代之以A_0B_1和$A_{n-1}B_2$的移动,那么我们所得到的(图22.6)就不是一种系带法.因此,我们将引进一种比系带法更广泛的对象,叫做箭号系统.在这个更大的门类中,所有的移动都将是合法的,而且经过有限次移动后,我们就到达了对
[197] 应于标准系带法的系统.

这些思想导致了下面的正式证明.

定理的证明 令$L = C_0 \to C_1 \to \cdots \to C_{2n+1}$为一种系带法(见图22.5),则存在某个$k$使得$C_k = A_n$.我们对$i = 0, 1, \cdots,$

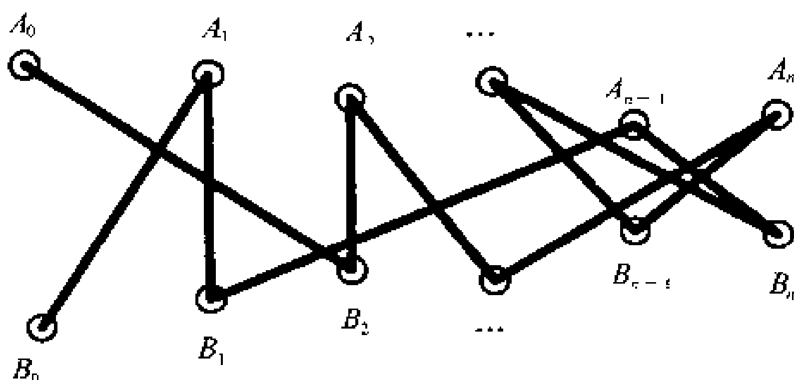


图 22.5 一种任意的系带法

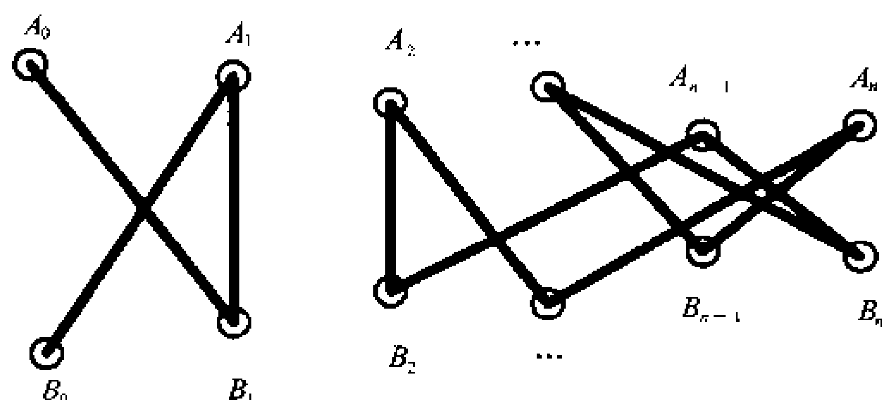


图 22.6 一次移动的结果

$k-1$,从 C_i 向 C_{i+1} 画箭号,而对 $i=k,k+1,\cdots,2n$,则从 C_{i+1} 向 C_i 画箭号(见图 22.7). 注意这个箭号系统满足下述性质:

(3) 每个箭号始于一个"A"鞋眼而终于一个"B"鞋眼,或者反之. 除了 A_0,B_0 和 A_n 外,所有的鞋眼都有一个箭号进来,一个箭号出去. 鞋眼 A_0 和 B_0 各有一个箭号出去而没有箭号进来. 鞋眼 A_n 有两个箭号进来而没有箭号出去.

我们定义一个箭号系统的长度就是其中箭号的长度之和. 虽然并不是每一个箭号系统都来自一种系带法,但如果是这样,那么它的长度就等于相应系带法的长度.

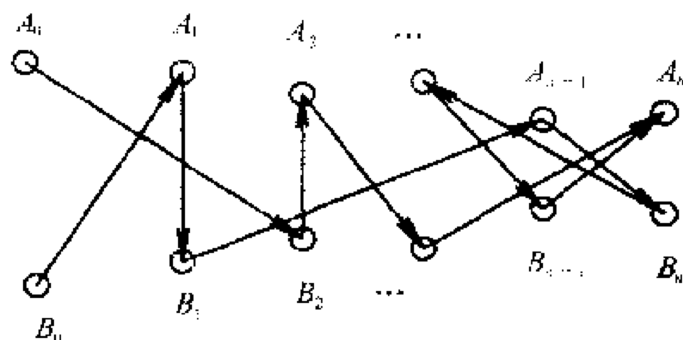


图 22.7　由图 22.5 中的系带法得来的一个箭号系统

让我们把从标准系带法得来的箭号系统称为**标准系统**. 我们证明它的长度小于其他任何满足(3)的箭号系统. 事实上,如果不是这样,那么就存在一个满足(3)的箭号系统 $\mathscr{A}$,它具有最小长度,但又不是标准系统. 如果对 $k = 0, \cdots, n-1$,箭号 $A_k \to B_{k+1}$ 和 $B_k \to A_{k+1}$ 都在 $\mathscr{A}$ 中,那么根据(3), $\mathscr{A}$ 就由这些箭号和箭号 $B_n \to A_n$ 组成. 这意味着 $\mathscr{A}$ 就是标准系统. 一个矛盾. 因此,我们可以取使得箭号 $A_k \to B_{k+1}$ 和 B_k①$\to A_{k+1}$ 中至少有一个不在 $\mathscr{A}$ 中的最小的 k. 我们可以假定是 $B_k \to A_{k+1}$ 不在 $\mathscr{A}$ 中. 于是,一定存在某个 $j \neq k+1$ 和某个 $l \neq k$,使得箭号 $B_k \to A_j$ 和 $B_l \to A_{k+1}$ 在 $\mathscr{A}$ 中. 因为当 $j \leqslant k$ 时任何终于 A_j 的箭号都始于 B_{j-1},而当 $l < k$ 时任何始于 B_l 的箭号都终于 A_{l+1},所以我们有 $j > k+1$ 和 $l > k$. 如果我们现在作一次移动,把箭号 $B_k \to A_j$ 和 $B_l \to A_{k+1}$ 代之以箭号 $B_k \to A_{k+1}$ 和 $B_l \to A_j$,则新得到的系统仍将满足(3). 但是根据(2)(令 $i = k+1$②),它的长度将小于 [198] $\mathscr{A}$ 的长度. 一个矛盾. 这就完成了证明.

① 原文误作 B_{k+1}. —— 译注

② 原文误作 $k-1$. —— 译注

第23章 太阳，月球与数学

当我听着那位博学的天文学家在演讲，
当那些证明和数字在我面前成列成行，
当他给我看那些图解和表格，并在上面做加减乘除和测量，
当我坐在这掌声阵阵的演讲厅里听那天文学家演讲，
我是那么快地陷入了无名的厌倦和迷惘，
我不得不站起身来悄悄地离场，
独自徘徊在潮湿而神妙莫测的夜色茫茫，
寂寥无声，我不时地抬头把星辰仰望．

——沃尔特·惠特曼[1]

也许我不该向惠特曼提出挑战，他已谢世，不能为自己辩护，但我不得不表示诧异，他为什么会被少数几个证明和数字弄得如此不耐烦．毕竟，那位博学的天文学家只不过是在做他该做的事，他在努力发现事物是怎样运行的．然而，我有一个感觉，即这首诗会让许多人感到宽慰，这些人一旦面对科学，就变得胆怯气馁，甚至会变得"厌倦和迷惘"．为听到沃尔特的诉说，宇宙在哥白尼(Copernicus)作出他关于太阳系的发现时也不知怎么一来变得不那么"神妙莫测"了．好了，对不起，惠特曼先

① 惠特曼(Walt Whitman，1819～1892)，著名美国诗人．这首名为《当我听着那位天文学家在演讲》(*When I Heard the Learn'd Astronomer*)的小诗发表于1865年，并不是他的佳作．——译注

生，但是我并不认为知道了天体力学的定律就必定会在某种感情水平上影响到一个人对夜空之美的反应能力．事实上，如果有什么影响的话，那情况正好相反．在我看来，另一位伟大的诗人说得很正确：真即美（而美即真）[①]——这是一件对我们大家[199]来说都很好的事，诗人和天文学家都一样．

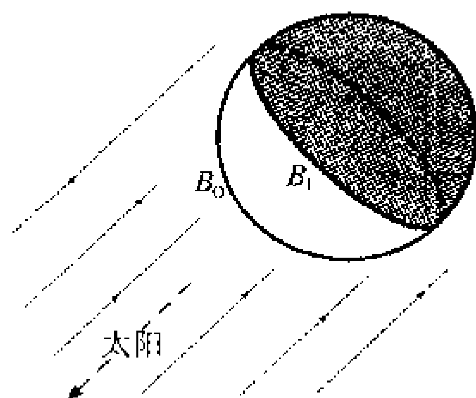
但不管怎么说，我要向沃尔特表示感谢，因为这首关于夜空的诗将成为我这次小宣讲的主题引文．而这次宣讲本身，除了展示一些基本的天文事实外，还打算说明几个准哲学观点，它们让我思索了好几年．第一个是数学就在我们周围这件事实．定理大量存在，只要我们睁大眼睛，留神注意，就能发现它们．第二个观点是关于数的避免使用．不仅在数学家中间，而且在一般人中间，存在着一种试图把什么东西都予以量化的倾向．但正如我过去曾说明的，有时候引进数会妨碍理解而不是帮助理解．因此，而且惠特曼也应当表示感激，我将不做"加减乘除和测量"．最后，或许我应该在我题目中的"数学"一词后面放上一个问号．读者将要判定，下面这些讨论是数学还是仅为"常识"．

月　　球

我们大家都曾经"不时地抬头仰望"，并看到一轮新月当空．一位当代的诗人可能把它的形状比做剪下来的指甲，但是在诗意较少的语言中，它可以被描写成一个以两条曲线为边界的区域，这两条边界我们称为它的外边界和内边界．于是我们的数学（?）问题就是要求描写这些边界．它们是圆弧，是椭圆弧，还是其他什么？在开始作了几次尝试，想引进坐标之类的东西而发觉不行之后，我画出了这幅规范的图．我们中许多人无疑在

① 这另一位伟大的诗人是指济兹（John Keats，1795～1821），英国诗人．这句话出自他的《希腊古瓮颂》（*Ode on a Grecian Urn*），原句是"Beauty is truth，truth beauty"，但这里引用时作"Truth *is* beauty（and beauty truth）"．——译注

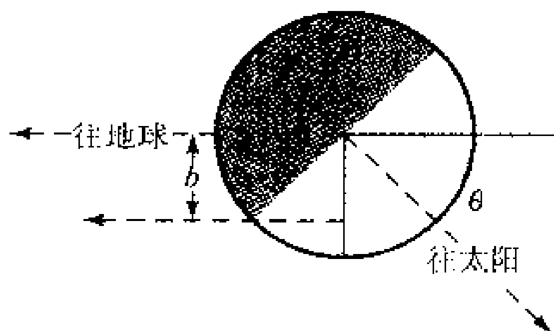
孩童时代看到过这幅图(我希望它不是那些让老沃尔特觉得如此令人不安的"图解"之一). 我们正从地球上直对着月球看去,用的或许是一副望远镜. 太阳在这图形所在平面背后的左下方. 那外边界和内边界记作 B_O 和 B_I.



现在这故事很简单. 我们考虑月球的两个半球面. 第一个是**可见**半球面,即我们在地球上可以看见的那个. 图中是我们直对着它所看到的情形. 第二个是**明亮**半球面,即被太阳照亮的那个. 这两个半球面的交,称为月亮,即我们平时所见的.

我们问题的答案现在清楚了. 外边界是一个半圆,它的直径就是月球的直径. 内边界是一个"半椭圆",它的长轴就是月球的直径. 这是因为如图所示,它就是月球"半赤道"的投影. B_I 的短轴 b 是什么呢? 为此我们还需要一幅图. [200]

这幅图所在的平面由地球、月球和太阳所确定. 设 r 是月球的半径,θ 是地球向月球发出的射线与月球向太阳发出的射



线之间的夹角，那么我们看到，那短轴 b 由 $b = r\cos\theta$[①]给出.

一个有趣的特殊情况就是当 θ 为直角时的半月. 在这种情况下，令 α 为地球向月球发出的射线与地球向太阳发出的射线之间的夹角. 伟大的希腊天文学家埃拉托斯特尼(Eratosthenes)利用这计算了月球与太阳到地球的距离之比. 具体地说，在一个月球和太阳都可以被看见的时刻，他测量了它们(从地球上看去的方位)之间的夹角 α. 于是到太阳与月球的距离之比就是 $\cos\alpha$. 这位非凡的人物还能够根据天文观察数据计算地球的直径. 结果他的估算值与我们现在的估算值相差不到五十英里!

我们要指出，我们问题的答案提供了一个关于"在自然界中实际存在的"椭圆的例子. 我想知道，除了行星的轨道之外，还有没有其他的椭圆例子?

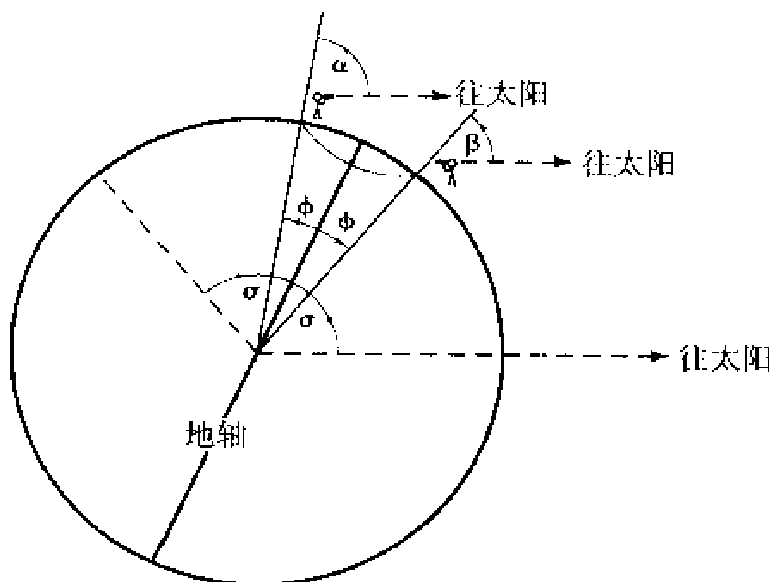
太　阳

既然我说过我要谈的是一些关于夜空的事，你可能会纳闷，怎样让太阳进入这幅画面呢? 在这一点上我们要离开大多数人的日常经验，除非你正巧是一位生活在北极圈以内的人. 在那种情况下，夏天有许多时候太阳一天24小时都在照耀着. 确实，我在最近一次去挪威的旅行中亲身体验了这件事，发现那情景真是令人叹为观止(我确信沃尔特·惠特曼也会有相同的反应). 但是它也使我想到一些问题，纵然相当简单，但仍是数学的. 或者说它们是数学的吗? 首先，我对自己说，在北极圈外太阳是东升西落，但当太阳既不升又不落时情况又怎样呢? 特别是，当太阳位于它的最低点，即最接近于地平线时，它在罗盘的哪一个方位上? 一个问题导致了另一个问题. 我于是又很想知道太阳"视轨道"的确切形状是什么. 当然，这个问题需要准确

① 事实上，这是那椭圆的短半轴. 但前面说这条内边界是"半椭圆"，因此它的短轴就是"全椭圆"短轴的一半了.——译注

表达. 让我们这样考虑. 你站在一个固定的地点,携带的设备是一副望远镜,可能还有一副高反射太阳镜. 每时每刻你都让你的望远镜正对着太阳,因此在一天当中,望远镜扫出了某种样子的一个锥面,并于24小时后回到了它原来的位置. 我们将把这锥面称为**太阳锥面**. 它是一种什么类型的锥面? 是圆的? 是椭圆的? 如果是圆的,它的中心在哪儿? 它的半径是什么?

下面这幅图给出了答案. 我们看到的是地球在太阳和地轴所决定的平面上(一个点和一条直线决定一个平面)的投影. 太阳远在水平方向上的无穷远处. 那位博学的天文学家,用♈表示(这是给沃尔特的一件特赠品,以表明天文学家们确实是人类),正站在纬度 ϕ 上,时值"午夜",即太阳最接近地平线的时刻;那副望远镜, →,正对着太阳,而太阳位于正北. 如图所示,天文学家所站处的球面法线给出了天顶的方向,而望远镜与天顶方向之间的夹角为 α. 在同一幅图上,到正午的时候,即12小时之后,太阳位于头顶稍稍偏南,而望远镜与天顶方向之间的夹角为 β. 人们可以试着粗略地画出这位天文学家在这一天其他时间的位置,但那只能起到把这幅图弄糟的作用. [201]



上面这幅图显示了在一个其中太阳固定不动的坐标系统中所实际发生的情况．现在我们的问题要把这种情况转移到这样一个系统中来描述，其中这位前哥白尼时代的天文学家固定不动而太阳在做着运动．显然，唯一有关系的是太阳方位与观察者所站处的地球法线之间的夹角．让我们假定这位观察者停留在原地并面向北方，就像我们图中上部的那两个人形．同样显然的是那太阳锥面完全由两个参数所决定：一个是纬度，或更方便于使用的余纬 ϕ[①]，即天文学家所站的地方；第二个，地球的“倾斜”角，即地轴与地球到太阳的连线所成的角 σ．现在容易看出，图中的角 α 就是 $\sigma+\phi$，而角 β 就是 $\sigma-\phi$，所以那锥面的生成元[②]与地轴所成的角就是 $(\alpha+\beta)/2=\sigma$．这幅图现在道明了真情．在太阳固定不动的描述系统中，那 ϕ － 锥面连同观察者围绕着地轴旋转．在观察者固定不动的描述系统中，是那 σ － 锥面做着围绕观察者的旋转，正如这图中虚线所表明的那样[③]．

① 即与纬度成余角的角．实际上，前面所说的（并在上图中标出的）纬度 ϕ 即余纬，显然余纬即球面坐标中的第二个坐标．作者从这里开始对纬度和余纬这两个概念加以区别．——译注

② 这里把一个锥面看成是由一条射线围绕着与其同一顶点的一条固定射线旋转而形成的，这条旋转的射线称作生成元．固定的射线称作中心射线，这两条射线间的夹角称作生成角．——译注

③ 虽然这幅图景应该说并不复杂，但作者如此叙述，理解起来恐怕并不是很顺当．为此，译者试作如下诠释：(1) 在上图中，让右边的那个小人举着望远镜向右方（太阳的方向）看，而且他的身子应该沿着该处的法线方向．这样，对他来说，就是在对着头顶上略偏南的太阳看．这是在太阳固定不动的系统中正午时的情况．左边的小人不变，这是这一系统中午夜时的情况．(2) 假设观察者固定不动，即不是地球在自转，而是太阳在围绕着地球“公转”．这时图上应只有一处有观察者，他停在原地随着太阳作自身旋转．我们设他始终站在左边的小人处，午夜时其状态即如上图左边的小人，正午时其状态应是我们如上修改过的图中右边小人关于地轴的镜射对称象．这是因为太阳绕地球“公转”的周期为24小时，且“公转”平面与地轴垂直，太阳12小时之前午夜时的方位与现在正午时的方位关于地轴呈镜射对称．于是太阳锥面在这幅图所在平面上的投影（两条同一顶点的射线之间的区域）的顶角为 $\alpha+\beta$，从而它的生成角就是 $(\alpha+\beta)/2=\sigma$．而它的顶点即观察者所站的位置．——译注

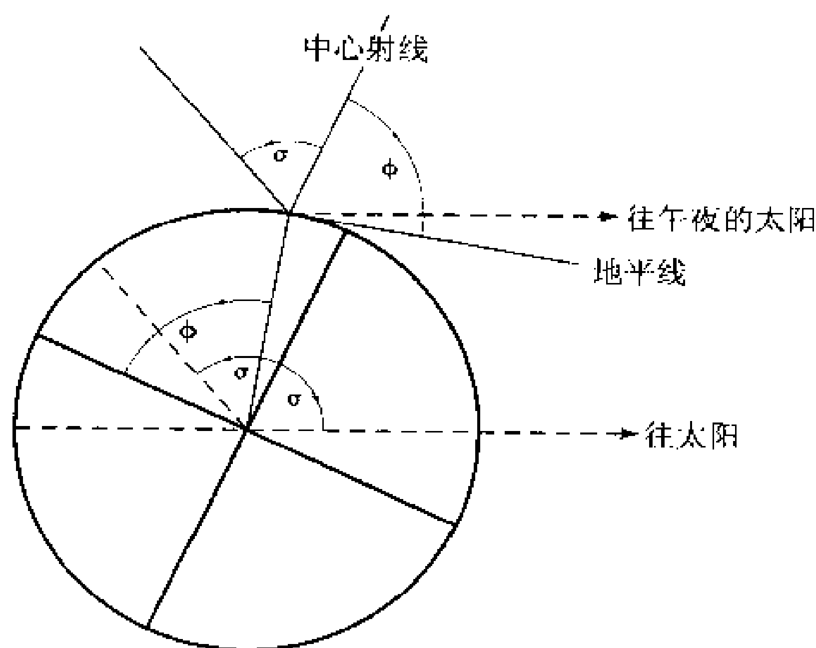
总而言之,观察者/天文学家所看到的太阳锥面,看上去就是一个以倾斜角 σ 为生成角旋转而形成的锥面,这个锥面的中心射线沿着地轴的方向,因此它相对于观察者的角是 ϕ,即观察者的余纬,或等价地说,这锥面的中心射线与地平线所成的角等于这观察者的纬度. 于是,这锥面的"半径"仅依赖于地球的倾斜程度,从而仅依赖于"季节",即地球相对于太阳的位置;但是它与观察者的位置(纬度)无关. 而这锥面的中心仅依赖于观察者的位置而与季节无关.

[202]

让我们把这稍微打扮一下,争取使它看上去更数学一些.

定理　太阳的视轨道是一个通过旋转而形成的锥面. 这锥面的中心由一条与从观察者向地平线发出的射线成角 ϕ 的射线给出,这里 ϕ 是观察者的纬度. 这锥面的生成角为 σ,这里 σ 是地球向着太阳的倾斜角.

请注意这午夜的太阳只不过是激发探索热情的动因. 这定理在一般情况下也是成立的,不管太阳是否走到地平线以下.



数　　学

或者说这是数学吗？我们似乎回答了我们所有的问题，而且这问题和答案看上去都是数学的，有椭圆、圆、角，等等；但是除了少数几幅图和一点儿常识外，我们又用到了什么？显然，我们根本没有给出过什么诸如对我们结论的正式证明之类的东西．仔细地用3维欧氏空间中的子集作为各种天文对象的模型来做这些事，大概是可以的．但在我看来，为这样一种方案付出努力是不值得的．事实上，一位具有足够兴趣来完成这样一个证明的读者，可能将不得不把那些符号回译成我们所讨论的直觉对象，以"看懂"这个答案．

还请注意在我们的结果中完全不存在数．这一点可能不那么十分明显．例如，人们可能会把纬度看作一个数，但这并不是必须的，而且是走上了歧途．纬度是一个角（地球上一个点的纬度是这点上的地球法线与地轴之间夹角的余角），而一个角不是一个数．一个角是从一个公共点出发的射线的一个有序对．同[203]样，一个椭圆的长轴和短轴也不是数，它们是线段．正如我以前曾指出的，数的概念从未在欧几里得的《几何原本》出现过，这里不需要它．

最后，回到我们的出发点；是的，沃尔特·惠特曼，很高兴被提醒想起在寂寥无声之中凝望夜空所体验到的敬畏感．但是我们当中也有一些人，他们从搞清楚一轮新月的确切形状这类似乎没有用处的活动中得到了一种满足感．问题在于我们周围到处都会产生令人惊奇的事物，但是它们以各种不同的形式来到我们面前，而如果我们能够敞开心扉迎接它们，那我们就是幸福[204]的．

第24章 没有数真好

避免使用数

作为一名数学家，一件较为令人沮丧的事就是，除了职业科学家以外，几乎没有人对这门学科是关于什么的或者我们在做什么事具有哪怕最模糊的理解. 随便找一个人来，问他数学的中心概念是什么，回答将很可能是**数**. 现在，我自己也不知道数学是关于什么的(后面对此还要详细谈)，但是我十分肯定地认为，这个中心概念不是数. 不过，存在着一个中心概念. 它就是**证明**. 下面讲到的例子打算用来说明，有时候本能地试图使用数的想法是怎样使证明变得更难的.

数　豆　子

考虑一个最幼稚水平上的问题，即判定两罐豆子中哪一罐豆子多. 太容易了，你说，只要数一数每个罐子中的豆子数目就行了. 不错，但是一个不知道怎样计数的人可能会发现下面这个更为快捷的解决方法. 同时从两个罐子中取豆子，每在一个罐子中取一颗豆子，在另一个罐子中也取一颗，如此成对地取. 当其中一个罐子被取空了的时候即停止.

205

这个寓言式的故事，(1) 说明了“并行计算”的优点，(2) 在一种非常初等的水平上说明了通过一一对应给出的集合基数的概念.

掺兑葡萄酒

稍微成熟一点的是这样一个众所周知的问题，其中说明了同样的思想. 给你一瓶红葡萄酒和一瓶白葡萄酒，要你取一调羹红葡萄酒，加到白葡萄酒中，彻底混匀. 然后取一调羹这种混合酒，加回到红葡萄酒中. **问题**：是白葡萄酒中的红葡萄酒多，还是红葡萄酒中的白葡萄酒多？（**警告**：这个问题可能有破坏人际关系的作用，有人会为他的错误答案同你发生激烈争论，例如，"你把纯红葡萄酒加到白葡萄酒中，而把一种混合酒加到红葡萄酒中，因此肯定是白葡萄酒中的红葡萄酒更多."）

现在人们当然可以用定量的方法把它解决. 令 R 和 W 为给出的红葡萄酒和白葡萄酒的体积（以某种单位计量），令 S 为那调羹的容积. 于是不难得出用 R，W 和 S 所表示的有关数量的式子. 问题是这完全是不必要的，而且，那彻底混匀的条件也是一种转移注意力的花招. 这里是另一个脚本. 你有一满瓶葡萄酒（不管红的白的）. 把其中的一些酒倒入大海，等 5 分钟，然后再用这稍稍被污染了的海水把酒瓶灌满. **问题**：是大海中的葡萄酒多，还是这葡萄酒中的（纯）海水多？

答案：这两个量相等. **理由**：在这次实验结束的时候，酒瓶中的海水**代替**了原本在那儿的酒. **练习**：失去的酒到哪儿去了？

这个例子的重要性在于它不仅避免了数，而且人们可以有理由认为，它也避免了任何类型的数学. 除了一些**常识**外，它什么也没有用到. 我们将在下面回到这个问题.

定性几何

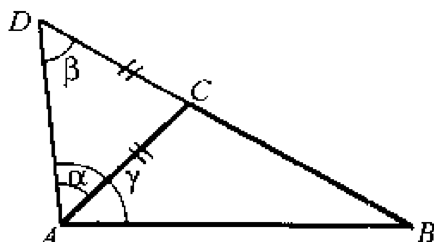
前面这些问题以及下面将要讲到的那些问题**都是定性**问题，与其相对的是定量问题. 定性问题处理的是像**多**、**少**、**高**、**相等**、**低**、**大**、**小**这样的概念，而定量问题就要把数联系到对象上去. 定性理论的一个杰出的例子就是欧几里得的几何，我领悟

到这一点后感到十分惊奇. 不要把这同我们现在所称的那种具有"欧氏度量"的欧氏几何混为一谈. 在欧几里得的《几何原本》中,没有度量,没有距离的概念,没有一个角的**大小**的概念. 不错,欧几里得进行了对象的加、减和比较,但是这些对象是线段和角本身,而不是它们的长度或大小. 例如,这里是三角不等式,即著名的第 20 号命题.

第 20 号命题 在任何三角形中,两条边接起来大于剩下的那条边.

下面是欧几里得的证明.

对给出的三角形 ABC,延长 BC 边到 D,使得 DC 等于 AC(这里以及其他地方,"等于"要读作"全等于"). [206]

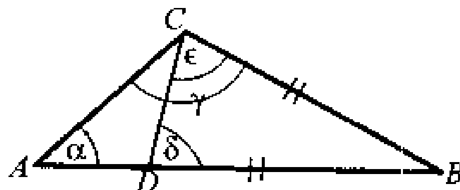


于是,由于三角形 ADC 是等腰的,角 α 和 β 就相等(第 5 号命题). 而且,α 小于(读作"被包含于")γ. 但是在三角形 ABD 中,γ 的对边是 $DC + CB$,而 β 的对边是 AB. 因此根据(前面已证明的)第 19 号命题,$DC + CB = AC + CB$ 大于 AB.

一切都是定性的! 不过让我们继续顺着原路往回走. 第 19 号命题是第 18 号命题的逆命题,它可从第 18 号命题直接得出.

第 18 号命题 在任何三角形中,大边对大角.

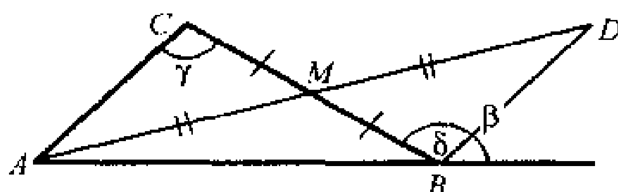
参见下图,我们假设 AB 比 BC 长.



在 AB 上取 D，使得 BD 等于 BC. 于是 CBD 是等腰三角形. 因此角 ϵ 和 δ 相等(第5号命题). 但是 ϵ 被包含于 γ，从而小于 γ. 而 δ 是三角形 ADC 的一个外角，因此它大于 α.

这个结论又用到了

第 16 号命题 一个三角形的外角大于任何一个内对角.



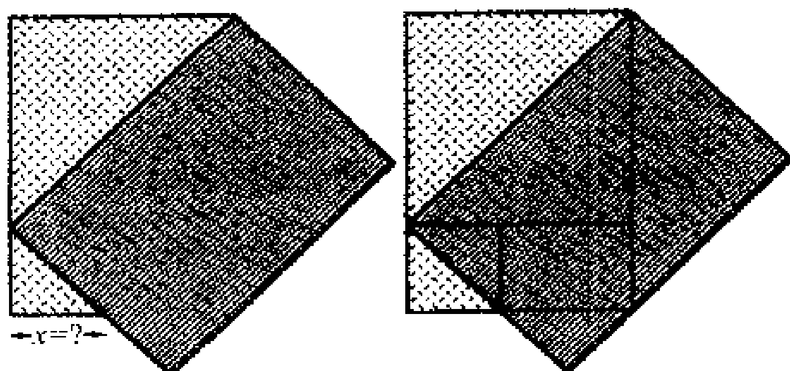
为证明 β 大于 γ，连接 A 与 BC 的中点 M，并延长线段 AM 到 D，使得 MD 等于 AM. 于是三角形 AMC 和 DMB 全等(边—角—边)，因此 γ 等于 δ，后者被包含于 β，从而小于 β. ■ [207]

评注 我们知道，欧几里得也知道，有一个强得多的事实，即这个外角是内对角之和(第32号命题). 为什么欧几里得会在这里甘心于这个弱得多的结果呢？这理由当然是他要在能不用平行公设就不用的情况下尽可能走得远一些. 用今天的语言，我们就说上面证明的这三个命题是绝对几何的定理，它们除了在欧氏几何中成立外，在椭圆几何和双曲几何中也同样成立. 我们又看到了一个例子，它显示了2000年前的几何那令人惊叹的成熟性.

叠 矩 形

有两个全等的矩形，如下页图左边所示，一个叠在另一个上面. **问题**：上面矩形覆盖了下面矩形的一半以上还是不到一半？

首先的一个冲动很可能就是把数带上这个舞台，即计算那两个未被覆盖的三角形的面积. 当然，人们可以这样做，但是，甚至在得到了多少有点杂乱的表达式后，也根本看不出来不等号应该朝哪个方向. **另一方面**，只要稍稍一想，就会想到画出如



图右边所示的那两条直线,并注意到有两对全等三角形. 事实上,人们甚至不一定非要有面积的概念才能提出这个问题. 我们只要同意这样的说法,如果一个区域"分块全等"于另一个区域的一个子集,那么后一个区域就大于前一个区域.

不管怎么说,数学是什么?

经过反复思考之后,我得出了结论:不存在像数学这样的东西. 让我解释. 请再看一下那个葡萄酒-海水问题. 如果人们把它适当地"解构"①,则所有的数学都消失了.

这里还有一个著名的问题,一点也不像是数学……或者说它是数学吗?

从镜子说起

有人问道,为什么镜子把左右对换,而不把上下颠倒?

这几乎不像一个数学中的问题,但如果它不是,那么它是一个什么门类的问题呢? 物理的? 光学的? 对这个问题有着一大批相当丰富的文献,但是我还得见到一个完全令人满意的讨论. 下面是我的看法.

[208]

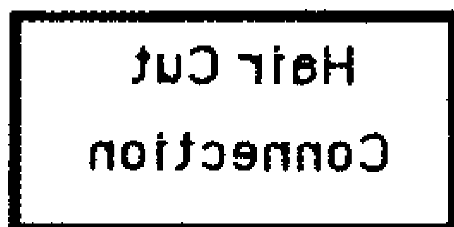
① 原文是 deconstruct,动词. 据查,相应的名词为 deconstruction,又"解构",或"拆析",文艺评论用语,指找出文本中的自身逻辑矛盾或自我拆解因素,从而摧毁文本在人们心目中的传统建构. 这里显然是借用.——译注

1. 这镜子是不相干的. 这种现象也出现在其他截然不同的背景下. 在一张薄纸上写下几个词,然后把它翻转过去背着你,并把它举起来对着光. 那些词还在那儿,但是次序与写法全部反了,而上下并没有颠倒. 这是纸的奇异特性吗? 当然不是. 这现象是由你和这个实验造成的,当把这张纸翻过来对着光的时候,你选择了让它绕着垂直轴而不是水平轴翻转的方式.

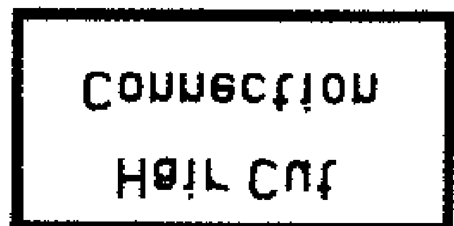
2. 我家附近的理发店有一个招牌,就画在它的玻璃窗上. 从街上看去,它是这样的:



但当理发师为我理发时,我从店内看去,它是这样的:



但不是像这样:



怎么回事? 还是我,即"实验者",导致了这些现象的发生,但在这种情况下,是我自己移动了位置而不是玻璃窗发生移动. 当从店内看这玻璃窗时,我面对的方向与我从街上看它时的方向正好相反,而我是通过把自己绕着一条垂直轴翻转来实现这种方向转换的. 如果我愿意,并且具有足够良好的肌肉协调能

力，我可能会在理发店内头朝下倒立着面对那玻璃窗．在这种情况下，那招牌的写法看起来就会像上面最后一幅图那样．

3．这种镜子现象被认为是有悖常理的，是因为镜子似乎把各侧面对换而不把上下颠倒．但这不是事实．我的镜子是在我卧室的北墙上，我的窗子是在西墙，看出去就是圣弗朗西斯科湾．当我向镜子里看时，镜子中的窗子也是在镜子中我卧室的西墙．显然，这镜子没有把东和西对换．那为什么我们被弄糊涂了呢？

4．然而，你会说，为什么镜子中那个被我正瞧着的家伙看上去是把手表戴在他的右腕上（而不是左踝上）．

问题是这些悖论是我们自己制造的．不同于东和西或上和下，**左**和**右**是在观看者的眼睛里．**向后**、**颠倒**也是这样．所有这些用黑体表示的词都是**相对于观察者**而定义的．而东、西、南、北、上、下是客观的．太阳总是从东边升起，不管我在什么地方，但它可能在我的左边或右边，前面或后面升起，这依赖于我的朝向．这里的谬误在于混淆了相对的与绝对的概念，混淆了"主观的"左－右与"客观的"东－西．我说我的镜射象把手表戴在他的右腕上而不是左踝上，是因为如果我要去镜子中代替"他"，我会通过向后转，脚保持站在地上而不是进入手倒立状态来做到这一点．我们人类，还有其他大多数动物，都是双侧对称的．这件事使我们觉得"转身"的唯一自然的方式是围着一条垂直轴而不是水平轴旋转，但事实上不一定非要按这种方式不可．

还是让我们回到所考虑的问题．上面这些讨论以什么方式可以被看作是数学的？我认为在形式上和内容上都存在着某种相似性．关于形式，就像在数学中那样，我们坚持认为我们所用的词语有准确的定义．关于内容，那关键词左和右，就是方位概念的日常表述，而方位的概念在数学中许多地方普遍存在．

这里最后的一个例子，甚至离人们通常认为是数学的领域更为遥远．我记得在一本儿童科学图书中读到过一件相当显然

的光学事实：物体在眼睛视网膜上的象是倒立的. 这本图书继续说道，大脑通过某种非凡的原因不明的技巧把它们设法回复到正立的状态. 这又是一例混淆主客观的谬误. 视网膜上的象对谁来说是倒立的? 对某个从外部窥视我眼睛内部的人. 说它在我看来是倒立的，就推出了这种荒唐的概念：我在看我自己的视网膜.

举这些例子的目的是阐明数学与非数学之间的灰色区域，但更重要的是，是表明如果我们没能足够严密地对词语进行分析的话，我们会怎样被词语捉弄并掉进词语陷阱. 在我看来，当我们试图定义数学时，我们就成了一种类似的词语陷阱的牺牲品.

寻找数学的一个定义就等于去捕捉鬼火. 举例来说，我们甚至没有一个关于"绿"的定义. 我的妻子坚持说我戴的这条绿领带是蓝的，而我们两人都没有办法能证明自己是正确的. 在文人雅士这个层次上，我想起许多年以前，在《纽约时报》(*New York Times*)的图书评论栏中，围绕"历史是什么"这个问题展开了一场旷日持久的辩论. 学者 A 说历史只不过是所发生的每一件事，但学者 B 坚持认为，不，历史只是那些发生的并使其他事情也发生的事情，或者大意是那样的话——争论无休无止. 我想这些事情例示了我所称的语言谬误. 我们发明了像"历史"、"数学"、"绿"这样的词，用来相互交流信息. 但是后来我们认为：因为这些词存在，它们一定对应着某种实际的外界实体；历史或数学的概念有着它们自己的某种存在物，或许就在上帝的心中，因此我们千方百计想发现这些事物**实际上**是什么，却忘记[210]了事实上正是我们，不是"他"或"她"，首先发明了它们.

数学是什么? 我的回答是不存在这样的**东西**. 对于某些事物，例如"上"、"下""左"、"右"，有定义是至关重要的. 对其他事物，像数学，寻找定义成了仅仅一种文字游戏. 我们不要在这上面浪费时间.

最后还有一个视点

这来自一位小孙女，她正在读二年级，她对数学这门课程的态度多少有点矛盾．照她的说法，她喜欢进位，但无法容忍借位——这大概说明了一切．通话完毕[①]． [211]

① 原文是 Over and out．系无线电通话用语．——译注 [212]

附录 1 一个奇特的 Nim 型游戏

一组对象，共有 mn 个，排成一个 $m \times n$ 的矩形阵列. 我们把位于第 i 行第 j 列的对象记为 (i, j). 第一位局中人 P_{I} 选取了一个对象 (i_1, j_1)，然后把所有满足 $i \geqslant i_1$ 和 $j \geqslant j_1$ 的对象 (i, j) 取走. 换句话说，如果 i 向上递增而 j 从左向右递增，那么 P_{I} 取走的是一个东北"象限". 现在第二位局中人 P_{II} 从余下的对象中捡起 (i_2, j_2)，并取走所有满足 $i \geqslant i_2, j \geqslant j_2$ 的 (i, j). 接着又轮到 P_{I} 走，如此继续下去，直到所有的对象都被取走. 走最后一着的局中人就是输家. 因此这个游戏的目的是迫使你的对手捡起 $(1,1)$.

这个游戏有一些平凡的特殊情况.

情况 A　在 $2 \times n$ $(m \times 2)$ 的游戏中，P_{I} 可通过选取 $(2, n)$ $((m, 2))$ 而获得胜利. 作了这个选取后，不管 P_{II} 怎么走，P_{I} 总是走得留下这样一个"局面"：第一行(列)的对象比第二行(列)多一个. 读者将很容易看出这种走法总是可行的，而且将导致胜利.

情况 B　在 $m \times m$ 的游戏中，P_{I} 可通过选取 $(2,2)$ 而获得胜利. 作了这样的选取后，他就不断采取"对称化"的走法. 每当 P_{II} 选择了 $(1, j)$，他就选择 $(j, 1)$，如此等等. 同样，容易看出[213] 这种走法将导致胜利.

上面所述的是两种仅有的已知一般必胜策略为如何的情况. 然而,使这个游戏令人感兴趣的是下面这个定理.

定理 *对所有的 m 和 n,这个游戏应该总是 P_{I} 赢.*

这个事实的证明完全是非构造性的,这种情况在对策论中经常出现,这就是一个典型. 虽然它确立了对 P_{I} 而言的一个必胜策略的存在性,但对找到这样一个策略绝对没有用处. 下面就是这个证明. 分两种可能.

情况1 P_{I} 有一个必胜策略,其中第一着是选取 (m,n).

情况2 如果 P_{I} 选取了 (m,n),结果输了,那么一定存在 P_{II} 的一个应着 (i_2,j_2),它使 P_{II} 获得了胜利. 这意味着,P_{II} 走了这着之后所造成的游戏局面对接下来要走的局中人(在这里是 P_{I})来说是一个必败局面. 然而问题在于,如果 P_{I} 先前就选择了 (i_2,j_2) 的话,他本来是可以把这个局面推给 P_{II} 的. 因此 (i_2,j_2) 对 P_{I} 来说是一个开局胜着,于是结论得证.

上述证明使人联想起对这样一个结论的著名证明:在像"连城"这样的游戏中,P_{II} 不可能有必胜策略(例如,对于尚未解决的五子棋对策,这个结论成立. 因为五子棋就是在一张 19×19 的棋盘上玩的"连城",但要求五子成一排). 这个证明如下. 假定对这个游戏 P_{II} 有一个必胜策略,那么就让 P_{I} 第一着随意走一步,并在心中自认他没有走过这步,在这之后,他按照 P_{II} 的那个假定存在的必胜策略行事. 这种做法总是可行的,除非走到某一步时这个策略要求他在他最初选择的格子里走子. 在这种情况下,他再次随意地走一着. 于是我们看到,这个对 P_{II} 假定存在的必胜策略对 P_{I} 同样有效,但是根据什么是取胜的定义,对局双方都不可能有必胜策略. 这就产生了一个矛盾. (然而,请注意,这里给出的对"连城"的证明是一种反证法,而我们那个定理的证明却是直接证法. 这就强调指出了这样一个有趣

的事实：一个非构造性的证明不一定非用反证法不可[①].)

可能让人们感兴趣的是，Nim 和我们这个游戏(是叫 Gnim? 还是叫 Gnome?)都是下面这个一般游戏类的特殊情况：令 S 为任意的偏序集．局中人的走法是选择 S 的某个元素，并把大于或等于这个元素的所有元素都取走，走最后一着的局中人就是输家．Nim 对应于 S 为有限个全序[②]集的和(不相交集合的并)的特殊情况．Gnim 就是 S 为两个全序集的笛卡儿积的情况．对 Gnim 应该总是 $P_{Ⅰ}$ 赢的证明，同样适用于任何具有最大元素的集合 S，例如任意多个全序集的笛卡儿积．不过，这个证明当然不适用于 Nim.

上面就是我所掌握的关于这个游戏的基本上全部的理论信息(我将在结尾给出一个更为特别的结果)．然而，在一台计算机的帮助下，已经获得了一些相当有趣的经验性结果．对 $3 \times n$ 的游戏就 $n \leqslant 100$ 的情况所作的全面分析揭示，在所有这些情况中，$P_{Ⅰ}$ 只有一个开局胜着．4×5 和 4×6 的游戏也是这样．当然，还有上面讨论过的情况 A 和情况 B.

---

① 译者以为，那个定理的证明事实上也用到了反证法，不过很隐蔽．情况 2 应该是情况 1 的补，作者用"$P_{Ⅰ}$ 选取了(m,n)，结果输了"这种说法表示情况 2，不免含混．严格地说，情况 2 应该是：(a)$P_{Ⅰ}$ 至少有一个必胜策略，但其中的第一着都不是取(m,n)，或(b)$P_{Ⅰ}$ 没有必胜策略，而 $P_{Ⅱ}$ 有必胜策略，或(c)$P_{Ⅰ}$ 和 $P_{Ⅱ}$ 都没有必胜策略．其中情况(c)可用策梅洛定理排除；情况(a)其实不必讨论，因为它等于说定理结论已成立；要排除的是情况(b)．但作者事实上还是把情况(a)和(b)放在一起讨论了．接下来的那句话"那么一定存在 $P_{Ⅱ}$ 的一个应着(i_2,j_2)，它使 $P_{Ⅱ}$ 获得了胜利"相当于：如果是情况(a)，而 $P_{Ⅰ}$ 错选了(m,n)，那么 $P_{Ⅱ}$ 有一个必胜策略，其中针对这种局面的走法是取(i_2,j_2)；如果是情况(b)，那么 $P_{Ⅱ}$ 的必胜策略中针对这种局面的走法是取(i_2,j_2)．下面的证明相当于：对于情况(a)，$P_{Ⅰ}$ 有一个必胜策略，其中的第一着是取(i_2,j_2)(一个多余的结论)；对于情况(b)，$P_{Ⅱ}$ 的这个必胜策略同样可以为 $P_{Ⅰ}$ 所拥有．这就产生了一个矛盾．其实，如果明确地使用反证法，即假设 $P_{Ⅰ}$ 没有必胜策略，从而 $P_{Ⅱ}$ 有一个必胜策略，证明将十分简单．——译注

② 关于"全序"的概念，可参见本译丛中《当代数学：为了人类心智的荣耀》的第 5 章．——译注

问题 是不是对所有的 m 和 n，这开局胜着都是唯一的？［214］（参见注记.）

下图对 $2 \leqslant n \leqslant 12$ 给出了 $3 \times n$ 游戏的开局胜着.

×	×	××	×××	×××
××	×	××	×××××	×××
××	×××	××××	×××××	××××××
2	3	4	5	6

××××
×××××××
×××××××
7

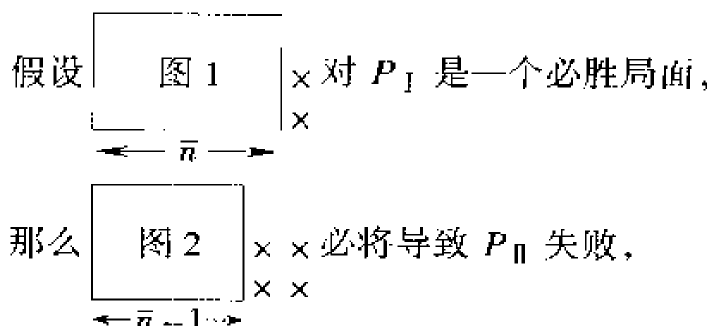
××××	××××××	×××××
××××	×××××××××	×××××
××××××××	×××××××××	××××××××××
8	9	10

××××××	××××××××
××××××	××××××××××××
×××××××××××	××××××××××××
11	12

显然，这里存在着两种走法，一种是第一口只在顶行中“啃”，另一种是第一口在上面两行中“啃”. 算下来大约有 58% 的走法属于这第二种，例如 $n = 3,4,6,8,10,11$ 的情况；而 42% 的走法属于第一种，如 $n = 2,5,7,9,12$. 大体上说，“啃”下来的长度看来随 n 而增加. 事实上，对所有的 $n \leqslant 170$，只存在一个反例. 对 $n = 87$，“啃”下来的尺寸是 37，而对 $n = 88$，“啃”下来的尺寸是 36（它们都是一口“啃”两行的）. 这样的现象诱导人们相信，要得到必胜策略的一个简单公式恐怕是十分困难的.

我们以一个猜想作为结束：除非 $m = 2$ 或 $n = 2$，第一着选取(m,n) 绝不会是胜着. 我们将对 $m = 3$ 的情况证明这个猜想，并以此作为对人们在进行非构造性论证时所用方法的又一个例示. 这里需要用反证法.

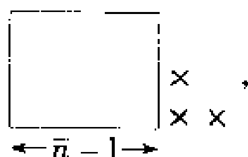
首先注意到在 3×3 的游戏中取$(3,3)$ 是败着(因为取$(2,2)$必胜). 现在假设对所有不大于 $\bar{n}$ 的 n，在 $3\times n$ 游戏中取$(3,n)$都是败着，于是考虑 $3\times(\bar{n}+1)$ 的情况. 证明最好用图形给出：



因此 P_{I} 一定有一种走法把图 2 走成一个必胜局面. 但是很显然，选择任何满足 $j \leqslant \bar{n}$ 的(i,j) 都不会给出一个必胜局面，因为这样将造成一个本来可以由 P_{II} 交给 P_{I} 的局面，这就同 P_{I} 取$(3,\bar{n}+1)$ 将导致胜利的假设发生矛盾. 于是可能的选择只有$(1,\bar{n}+1)$ 或$(2,\bar{n}+1)$[①]，但是取$(1,\bar{n}+1)$ 将造成



[215] 根据归纳法假设，这是一个败着. 而如果取$(2,\bar{n}+1)$，则造成



① 这里及下面的$(1,\bar{n}+1)$ 和$(2,\bar{n}+1)$，原文分别都错作$(\bar{n},1)$ 和$(n,2)$. ——译注

接下来 P_{II} 可以走(3,1)，造成

```
┌──────┐×
└──────┘× ×，
```

这将导致 P_{I} 失败，因为它就是情况A中 $2\times n$ 游戏的局面. 这就完成了证明.

像这种特殊的结果，人们可以用类似的方法证出任意多个. 例如，对于 $n>4$，选取 $(2,n-1)$ 绝不会是胜着；对于 $n>5$，选取 $(3,n-1)$ 也绝不会是胜着. 而且可以推测，存在一些一般的不等式，它们表明对于大的矩形阵列，"啃"下来的那口不会太小. 我估计要找出详细的必胜策略恐怕是没有希望的.

最后，让我简单介绍一些推广. 首先是允许 m 或 n 或它们同时为无穷大. 不过，这些游戏相当平凡，因为(a) $1\times\infty$ 应该总是 P_{I} 赢(很平凡)，(b) $2\times\infty$ 应该总是 P_{II} 赢(这为读者提供了一道美妙的练习题)，(c) $m\times\infty$，$2<m\leqslant\infty$，应该总是 P_{I} 赢，因为他可以选择(3,1)[①]，给 P_{II} 造成一个 $2\times\infty$. 更让人感兴趣的是高维的游戏，例如三维空间中的 $m\times n\times r$. 当然，任何这样的有限对策都可以在一段有限的时间内决出胜负，因为充其量不过是枚举所有可能的策略. 在我看来，真正的挑战是像 $3\times3\times\infty$ 甚至 $\infty\times\infty\times\infty$ 这样的游戏. ($2\times m\times\infty$ 应该总是 P_{I} 赢. 为什么?) 这些游戏属于一个十分有趣的游戏类，其中的游戏具有这样的性质：虽然游戏的每一局在走了有限着以后就会结束，但一局的可能长度却根本不存在上界(例如，就像国际象棋那样). 特别是，我不知道有什么方法可以为计算机编一个程序，来解决像 $3\times3\times\infty$ 是应该 P_{I} 总是赢还是应该 P_{II} 总是赢这样的问题.

注记 对Gnim的一个描述出现在《科学美国人》上由马丁·加德纳主持的专栏中(*Scientific American*, January(1973)，

---

① 原文误作(2,1). ——译注

p.110—p.111). 作为对这篇文章的回应,贝尔实验室的汤普森(K.Thompson)和麻省理工学院的比勒(M.Beeler)用计算机发现,存在不止一个开局胜着的游戏. 已知最小的例子为 8×10. 另外,人们已经知道,舒(F.Schuh)在一篇题为《除数的游戏》(Game of Divisors[①])的文章中描述了一个同构于这个游戏的数值游戏[②](见《数学新杂志》(*Nieuw Tijdschrift voor Wiskunde*) [216] 39(1952),p.299—p.304).

---

① 原文误作 Game of Devisors,现更正. 事实上,这篇文章系荷兰语,题目为 Spel van Delers. —— 译注

② 这个游戏也很容易描述:任取两个不同的素数 p_1 和 p_2,再任取两个正整数 m 和 n,算出 $k=p_1^m p_2^n$. 于是对局双方轮流选择 k 的因数,每次一个,但每次都不能选已选因数的倍数. 被迫选择 1 的就是输家. 例如,令 $p_1=2$, $p_2=3$, $m=n=2$,则 $k=36$. 对局过程可以是:6,9,4,3,2,1. 容易证明,这个游戏同构于 $(m+1)\times(n+1)$ 的"大嘴巴".—— 译注

附录 2　再说吉普车——吉普越多越俭朴

引　　言

1947 年，法恩解决了现已十分著名的吉普车问题[1]．这个问题讲的是一辆加足燃料时能行进距离 d 的吉普车，但是却要求它穿越一片宽度大于 d（比方说是 $2d$）的沙漠．它可以用以下方法完成这件事：把燃料从它的大本营运送到沿途的不同地点建立燃料储藏点，以使它在向更远的地方行进时能够得到燃料补充．它必须以尽可能小的燃料量穿过这片沙漠．

这里我们的目的首先是给出这个吉普车问题的解的一个非常简短的推导，它用到了巴拿赫的一个定理（也是十分著名）[2]．其次，我们考虑每天——比方说在一星期中——都需要派一辆吉普车穿越沙漠的情况．当然，既然对一辆吉普车已经找到了最好的行进过程，人们只要把这个过程重复七次即可．然而，我们将证明对七辆吉普车的情况存在一种更为节俭的行进过程，而且一般地说，人们派遣穿越沙漠的吉普车越多，平均每辆吉普车的燃料消耗就越低．这种现象是经济学家所谓的"规模报酬递增"（increasing returns to scale）①的一个例子，这是一个在经济学上有着某种重要性的课题．（这也是这篇文章又取了一 [217]

① 所谓"规模报酬递增"，指的是某一产品或行业的净收益的增长速度超过其生产规模的扩大速度的现象或状态．——译注

个题目的原因，对此我在这里顺便表示歉意.)

问　　题

为了把这个问题形式化，让我们假设这辆吉普车从原点出发，沿着正 x 轴前进. 我们把这辆吉普车所能承载的最大燃料量取作燃料的单位，并把这个单位称作一**满箱**(load). 于是距离的单位就取为这辆吉普车用一满箱燃料所能行进的距离.

图 A 2.1 是一幅表示了一个典型的吉普车行程的简图.

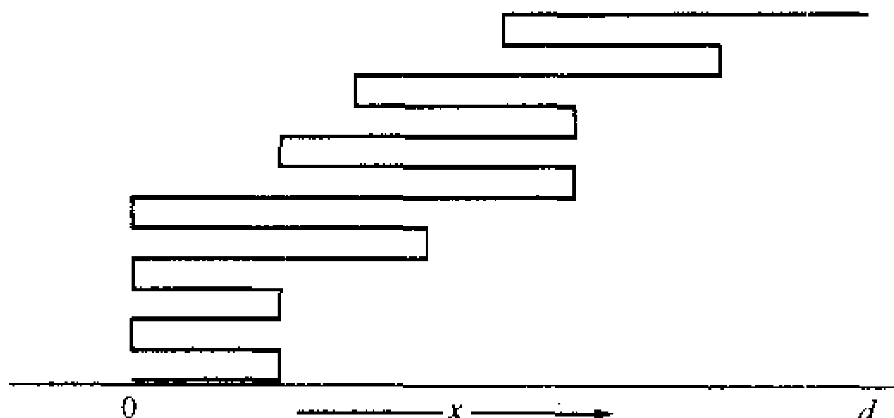


图 A 2.1

这条前后摆动的路径代表了这辆吉普车的行进历程. 当然，这条路径事实上完全躺在 x 轴上. 把它在垂直方向上拉开只是为了使它看得见. 由于我们对单位量的取法，这条路径的长度准确地等于燃料的消耗量. 在图中，吉普车到达了与原点距离为 d 的一个点. 本来那个吉普车问题要求的是能让吉普车到达这个点的最小燃料消耗量(从而最短的路径长度). 在某种程度上更为方便的做法是把这个问题转个向(比如参见[3])，求得一个表示函数 $d(f)$ 的公式，其中 $d(f)$ 给出了这辆吉普车用 f 满箱燃料能够到达的最远点. 我们的第一个目标是证明公式

$$d(f) = 1 + \frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{2f-1}. \tag{1}$$

我们解答中的关键思想是利用巴拿赫的一个公式，这个公式是关于一维空间中曲线的路径长度的.（一维空间中的路径长度通常称作**全变差**；而我们偏爱用这个几何学术语是因为它在目前的背景下更能帮助想像.）为了利用巴拿赫的公式，我们对区间$[0,d]$上的每个点 x 定义**价函数**(valance)[①] $n(x)$，它表示那辆吉普车在行进过程中到过点 x 的次数. 图 A 2.2 给出了对应于图 A 2.1 中所描绘的吉普车行程的价函数图象. 严格地说，如果我们允许有十分广义的路径，那么 $n(x)$ 在一些点上可能为无穷大. 但这不会对巴拿赫的公式产生影响. 这个公式是

$$\text{路径全长} = \int_0^d n(x)\mathrm{d}x.$$

[218]

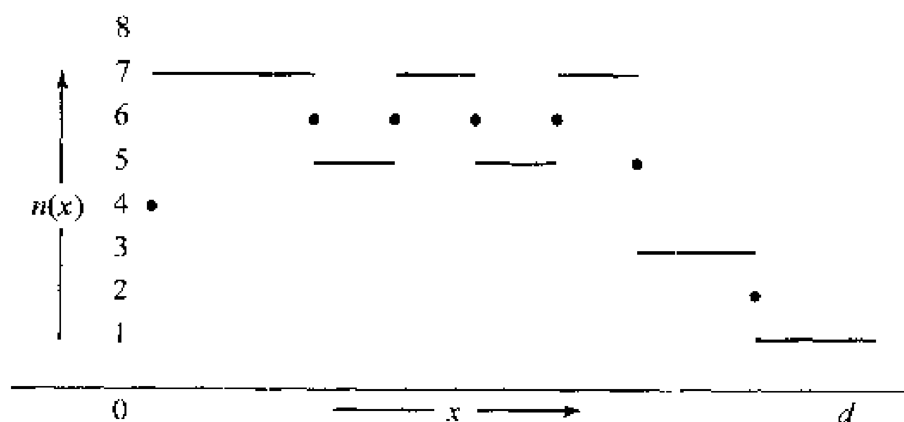


图 A 2.2

当然，巴拿赫并没有关注过吉普车. 他考虑的是一个连续的实值映射 $x=\phi(t)$，而 $n(x)$ 只不过是映射到点 x 的点的个数，或者说集合 $\phi^{-1}(x)$ 的元素个数. 我们同样可以这样看这个问题，只要我们认为图 A 2.1 中那条来回摆动的线是把吉普车位置作为——比方说作为时间的一个函数而描绘下来的图象.

---

① 在文献中一般又称"巴拿赫的指示函数"，如《实变函数论(上册)》(И·П·那汤松著，徐瑞云译，陈建功校，人民教育出版社 1958 年第 2 版）的 254 页.——译注

对于"合理"的路径,巴拿赫的公式是显然的. 一辆吉普车走出的一条合理路径只有有限多个调头点(一辆能走出一条不合理路径的吉普车实在是很了不起). 为了证明这个公式,把 $[0,d]$ 划分为集合 $X_1,X_2,\cdots$,其中

$$X_k=\{x\mid n(x)=k\}.$$

由于路径的合理性,只存在有限多个非空集合 X_k,它们每一个都是一些不相交区间的并.(它们中有一些可能由一些孤立点组成. 在这个吉普车运行的实例中,这种情况当 k 为偶数时发生. 你知道为什么吗?)在 X_k 的每个区间上,恰好躺着吉普车路径的 k 个区间. 因此

$$\text{路径全长}=\sum_{k=1}^{\infty}k\cdot(X_k\text{ 的长度}).$$

但是右边的项正是 $\int_0^d n(x)\mathrm{d}x$ 的定义.(注意这是勒贝格意义下而不是黎曼意义下的积分定义!当然,对连续函数,特别是对合理的函数来说,这两个概念是等价的.)我们指出,巴拿赫的公式对合理的和不合理的路径同样成立(虽然一般的证明相当复杂),因此我们对这个吉普车问题的解答对合理的和不合理的吉普车同样成立.

关于一辆吉普车的解

现在我们回到计算 $d(f)$ 的问题,并暂时假设 f 是一个整数. 我们希望确定这辆吉普车用 f 满箱燃料能走多远. 对于任[219]意的吉普车行程,我们在区间 $[0,d]$ 上定义点列 $x_0,x_1,\cdots,x_f$,其中 $x_f=0,x_0=d$,而一般地,x_k 是在其右边的路径总长(从而燃料消耗量)恰好为 k 个单位的点. 显然,这些 x_k 点形成了一个严格递减的序列,而且在 x_{k+1} 和 x_k 之间恰好有一个单位的路径长度. 我们所需要的基本观察结果是下面的

引理 1 如果 $x<x_k$,那么

$$n(x) \geqslant 2k+1. \tag{2}$$

证明 既然 x 在 x_k 的左边，那么这辆吉普车在 x 右边消耗的燃料一定多于 k 满箱. 既然这辆吉普车每次只能承载一满箱的燃料，那么它一定至少从左边经过 x 点 $k+1$ 次. 但是凡从左边经过两次，其间必定要从右边经过一次，因此一定要从右边至少经过 k 次. 于是这辆吉普车一定到过 x 点至少 $2k+1$ 次. 这正是这条引理的结论.

现在我们把巴拿赫的公式同这个结果结合起来，得到

$$\begin{aligned} 1 &= x_{k+1} \text{ 与 } x_k \text{ 之间的路径长度} \\ &= \int_{x_{k+1}}^{x_k} n(x)\mathrm{d}x \geqslant (2k-1)(x_k - x_{k+1}). \end{aligned}$$

因此 $x_k - x_{k+1} \leqslant 1/(2k+1)$. 从 0 到 $f-1$ 求和，即得

$$\begin{aligned} \sum_{k=0}^{f-1}(x_k - x_{k+1}) &= x_0 - x_f \\ &= d \leqslant 1 + \frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{2f-1}, \end{aligned} \tag{3}$$

这就给出了 $d(f)$ 的一个上界. 剩下来只要证明这个上界是可以达到的，这件事用数学归纳法很容易完成. 当 $f=1$ 时有关的式子显然是正确的. 假设我们现在可以使用 $f+1$ 满箱的燃料. 于是令这辆吉普车把这 $f+1$ 满箱燃料带着运到点 $1/(2f+1)$. 这就需要前行 $f+1$ 次而返行 f 次，从而要作 $2f+1$ 次距离为 $1/(2f+1)$ 的运行. 因此消耗了整整一满箱的燃料，留下 f 满箱储藏在点 $1/(2f+1)$. 根据归纳法假设，从这一点以后(1) 式成立. 这就完成了证明.

对于 f 不是整数的情况，同一类型的证明给出

$$d(f) = 1 + \frac{1}{3} + \cdots + \frac{1}{2[f]-1} + \frac{\{f\}}{2[f]+1}, \tag{4}$$

其中 $[f]$ 和 $\{f\}$ 分别是 f 的整数部分和小数部分. 换句话说，人们只要在 f 的整数值之间作线性插值. 图 A 2.3 描绘了 $d(f)$ 的

图象.

既然奇调和级数是发散的,那么就可以推出任意大小的沙漠都能被穿越.

在找到了刚才给出的解答之后,我才知道在[4]中也给出了一个类似的解答. 我们这里用了巴拿赫公式,看来这使得证明在某种程度上更为紧凑,推理也更为清晰. 而且,我们不需要事先假设只有有限多个储藏点. 至少可以允许最优解包含无穷多个储藏点,甚至包含连续地散布在一路上的某种黏稠性燃料. [220] 人们无疑能用测度论的术语表述出一个非常广义的问题. 然而,上述论证表明——还是多亏了巴拿赫的公式——这种更为广义的行为不可能对燃料的消耗情况有任何改善.

在离开这辆孤独的吉普车之前,我们来考虑要求这辆吉普车穿过沙漠然后返回的情况. 对于这种情况,有关的证明与前面相同,除了(2)变成

$$n(x) \geqslant 2k + 2. \tag{5}$$

令 $\bar{d}$ 为可能最长的来回路程(即到最远点的距离的两倍),我们得到了甚至更为简单的公式

$$\bar{d}(f) = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{f}, \tag{6}$$

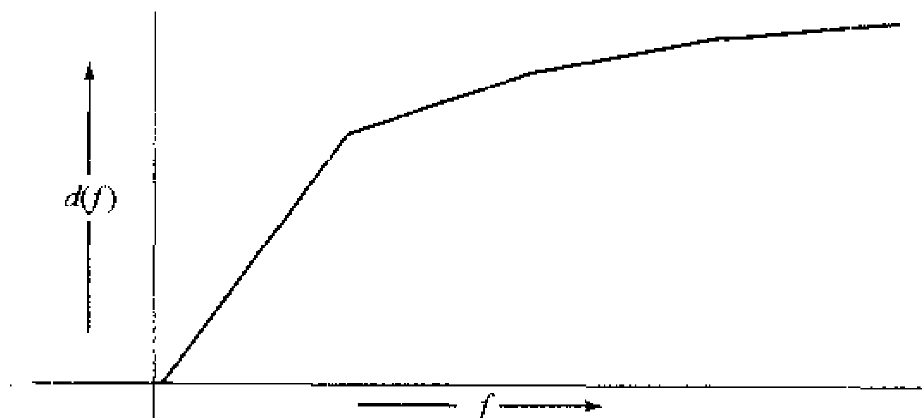


图 A 2.3

与前面一样，这里的等号是成立的．我们指出一个熟知的事实：一张来回票比两张单程票要便宜得多．事实上，把(1)同(6)比较，我们有

$$d(f)-\frac{1}{2}\bar{d}(f)=1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\cdots-\frac{1}{2f},$$

但是它以 $\sum_{n=1}^{\infty}(-1)^{n+1}(1/n)=\ln 2$ 为界．这样，对于很长的距离来说，一次来回旅行相对于一次单程旅行在燃料费用上的增加是可以忽略不计的．

多辆吉普车

现在我们转向 m 辆吉普车的问题，我们将计算函数 $d_m(f)$，它给出了这些吉普车全体能够到达的最远地点，条件是它们有 f 满箱燃料可以共享，换句话说，这些吉普车所走路径的长度之和不能超过 f．

我们从定义点 x_k 开始．就像一辆吉普车的情况，x_k 是在其右边全体吉普车所走路径长度之和为 k 的点．然而，与引理1对等的引理却稍稍复杂一些．

引理 2 对于 $[0,d]$ 中的任何点 x，$n(x)\geqslant m$．如果 $x<x_{m+r}(r\geqslant 0)$，那么

$$n(x)\geqslant m+2r+2. \tag{7}$$

[221]

证明 第一个结论只是相当于所有 m 辆吉普车都要到达 d 点这个事实．关于不等式(7)，我们知道必须要有多于 $m+r$ 满箱的燃料运到比 x 点更远的地方，所以这一点一定至少从左边被经过 $m+r+1$ 次．但是既然只有 m 辆吉普车，那么要做到从左边经过 $m+r+1$ 次，就一定要从右边经过至少 $r+1$ 次，于是总共要经过 $m+2r+2$ 次．

接下来我们计算 $d_m(f)$．注意对 $f\leqslant m$，这个问题是平凡

的，而且 $d_m(f) = m/f$（为什么？），所以我们假设 $f = m + s$，其中 $s > 0$. 在仍然把 s 限制为整数的前提下，我们宣称有

$$\begin{aligned} d_m(f) &= 1 + \frac{1}{m+2} + \frac{1}{m+4} + \cdots + \frac{1}{m+2s} \\ &= 1 + \frac{1}{m+2} + \frac{1}{m+4} + \cdots + \frac{1}{2f-m}. \end{aligned} \tag{8}$$

由于从 x_{k+1} 到 x_k 的路径长度为 1，

$$1 = \int_{x_{k+1}}^{x_k} n(x)\mathrm{d}x \geqslant m(x_k - x_{k+1}), \quad 对\ k < m,$$

以及

$$1 = \int_{x_{m+r+1}}^{x_{m+r}} n(x)\mathrm{d}x \geqslant (m+2r+2)(x_{m+r} - x_{m+r+1}).$$

因此，对 $k < m$，有 $x_k - x_{k+1} \leqslant 1/m$，而对 $r \geqslant 0$，有

$$x_{m+r} - x_{m+r+1} \leqslant \frac{1}{m+2r+2}.$$

对 $0 \leqslant k \leqslant m-1$ 以及 $0 \leqslant r \leqslant s-1$ 求和，就把式(8)作为不等式给出. 还是用归纳法来证明这等号是可以成立的. 假设这等式对 $m+s$ 满箱燃料是正确的，我们看到可以把 $m+s+1$ 满箱燃料带着运行到点 $1/(m+2s+2)$，方法是让一辆吉普车跑 $s+1$ 个来回，然后让所有 m 辆吉普车向这点作单程运行. 这将用掉一满箱燃料，因此在这点储存了 $m+s$ 满箱的燃料，从而归纳推理得以通过. 执行这个最优运行的方式有多种多样. 有一种方式是让其中一辆吉普车来完成建立各个储藏点的全部工作，其他吉普车只要在一路上补充燃料，头也不回地向前行进. 然而，规定所有的吉普车要走同样的路径却是不可能的，因为如果所有的吉普车都要循同样的路线行进，那么函数 $n(x)$ 在所有的点上都得被 m 整除，而在一次最优的运行中显然不会这样. 我们在后面将回到这一点.

为证明关于报酬递增的结果，我们要把一辆吉普车用一满箱燃料能走的距离同 m 辆吉普车用 mf 满箱燃料能走的距离作

比较. 为此,我们就 $s = m(f-1)$ 的情况把式(8) 改写如下:

$$d_m(mf) = 1 + \left(\frac{1}{m+2} + \cdots + \frac{1}{3m}\right) + \left(\frac{1}{3m+2} + \cdots + \frac{1}{5m}\right) + \cdots + \left(\frac{1}{(2f-3)m+2} + \cdots + \frac{1}{(2f-1)m}\right), \quad (9)$$

其中每个括号项包括 m 个被加项,一共有 f 个括号项. 我们也可以把式(3) 改写为

$$d(f) = 1 + \left(\frac{1}{3m} + \frac{1}{3m} + \cdots + \frac{1}{3m}\right) + \left(\frac{1}{5m} + \cdots + \frac{1}{5m}\right) + \cdots + \left(\frac{1}{(2f-1)m} + \cdots + \frac{1}{(2f-1)m}\right), \quad (10)$$

其中在每个括号项中也是有 m 个被加项. 对式(9) 和式(10) 就括号项进行逐项比较,就显示了使用多辆吉普车的优点. 对于 $m = 2$ 的特殊情况,我们得到

$$\begin{aligned} & d_2(2f) - d(f) \\ = & \left(\frac{1}{4} - \frac{1}{6}\right) + \left(\frac{1}{8} - \frac{1}{10}\right) + \cdots + \left(\frac{1}{4f-4} - \frac{1}{4f-2}\right) \\ = & \frac{1}{2}\left(\frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \frac{1}{5} + \cdots + \frac{1}{2f-2} - \frac{1}{2f-1}\right). \end{aligned}$$

于是行程越长,节约越大. 另一方面,由于上面括号内的项是一个收敛级数的部分和,结果是当行程越来越长时,节约下来的总量仍然保持有界.

对于来回旅行问题,式子还是比较简单. 令 $\tilde{d}_m$ 是 m 辆吉普车用 f 满箱燃料所能走的来回路程,则有

$$\tilde{d}_m(mf) = 1 + \frac{1}{m+1} + \frac{1}{m+2} + \cdots + \frac{1}{mf}. \quad (11)$$

因此,仍然是吉普车越多,它们用同样数量燃料所能走到的地方就越远,或者回到原来的问题,走到一个事先指定的地点平均每辆吉普车所需要的燃料就越少. 但是,增加吉普车数目而得到的利益有一个极限,因为当 m 趋近于无穷大时,一个常规的计

算（与$\int dx/x$进行比较）表明

$$\lim_{n\to\infty}\tilde{d}_m(mf) = 1 + \ln f. \tag{12}$$

所以不管有多少辆吉普车，平均每辆吉普车用 f 满箱燃料不可能走得比 $1 + \ln f$ 更远．用原来问题的话说，我们从这极限式中得到

$$f = e^{d-1}, \tag{13}$$

因此所需要的燃料量随着那沙漠宽度的增加而呈指数式上升，这正如人们可能已经预料到的．

这个多辆吉普车的来回旅行问题可以转述成另一种形式．本来是考虑 m 辆吉普车，每辆都要走一个来回，其实人们可以考虑单单一辆吉普车，它必须在接连的几天中，比方说在 m 天中，每天走一个来回．这里的解答同 m 辆吉普车问题的解答一样，从而燃料节约量也一样．事实上，这辆吉普车在第一天中可以把后面几天所需要的所有储藏点都建立好，人们应该很容易[223]想到这一点．

一些总结性的评注和问题

上面最后一段文字引出了一个有趣的问题．假设人们并不着意于把沙漠来回穿越 m 次，而是决定投身于穿越沙漠的营运业务，并且计划在未来一个长短不确定的时期内每天来回运行一次．那么人们应该采用一种什么样的例行程序呢？正如我们刚才看到的，每天都照例单走一个来回将是不经济的．同样，把计划 m 天走 m 个来回的工作程序照例重复地执行，其效益低于计划 $m+r$ 天走 $m+r$ 个来回的工作程序，所以这类周期性的程序根本不可能是最优的．另一方面，如果人们打算考虑非周期性的程序，那么就再也搞不清楚一个最优程序是什么意思了．无论怎样，要求最好的“稳态”例行程序的问题看来根本不会有准确的解答，所以在实践中人们将不得不满足于一个“几乎最

优"的例行程序.

人们还有许多吉普车问题可以考虑. 赫尔默(Helmer)[5] 考虑了一些相当复杂的情形,其中人们被允许建立的储藏点个数是有限制的. 要感受一下这种类型的问题,读者可以考察要穿越的沙漠宽度为 2 而途中只能有 3 个储藏点的问题.

一个表面上看来很简单的问题是在沙漠两头都有燃料供应的来回旅行问题,但我必须不好意思地坦白承认,我还不能找出解答. 不难看出,在这种情况下人们至少可以做到与两辆吉普车作单程旅行的情况一样的程度,但是还有可能做得更好些. 与许多最优化问题一样,这里的困难在于看来没有任何简单的方法可以判定一个给出的解是不是最优的. 我把这个问题提出来,作为对吉普车问题专家的一个挑战.

我用一些历史评注作为结束. 法恩的解答发表后不久,菲普斯(Phipps)[3] 得到了同样的结果,他的方法是证明这个单辆吉普车问题等价于一个护运车队问题——车队中的吉普车一路同行,并用其中一些车辆为其他车辆补充燃料. 对这个护运车队问题的解答非常简单,但是这个问题等价于原来问题的证明似乎不太完整. 不过,利用这个等价性,人们也能容易地得到这里给出的关于收益递增的结果. 最后,许多人有一种感觉,认为这个吉普车问题能够用动态规划中的函数方程法解决. 事实上,这个问题作为一道习题出现在贝尔曼(Bellman)的书中[6],但那里没有给出答案,而我也不知道怎样用这种方法解决这个问题.

补　　遗

有人向我指出,如果人们接受菲普斯的护运车队等价性,那么就可以用上动态规划方法(比如参见[7]). 然而,对护运车队问题来说,无论怎么看解答都几乎是显然的. 设想有 f 辆吉普车动身出发. 既然除了最后那辆外,所有的吉普车都必须返回,那

么我们就假定其中 $f-1$ 辆以 2 的比率消耗燃料，而最后一辆，即那辆"红白蓝"[①]吉普车，将最终完成穿越沙漠的任务．它以 1 的比率消耗燃料．这样，这个护运车队出发时以 $2f-1$ 的比率消耗燃料，并维持这个比率直到一满箱燃料被消耗掉，此后第一辆吉普车就被遗弃．余下的车辆继续前进，并以 $2f-3$ 的比率消耗燃[224]料，如此等等．

我还应该指出，对于赫尔默的那个储藏点个数被指定的变化问题，动态规划看来确实是一个适用的工具．

参考文献

1. N. J. Fine, The jeep problem, *American Mathematical Monthly* 54 (1947), 24—31.
2. S. Banach, Sur les lignes rectifiables et les surfaces dont l'aire est finie, *Fundamenta Mathematicae* Ⅶ(1925), 225—236.
3. C. G. Phipps, The jeep problem, A more general solution, *American Mathematical Monthly* 54(1947), 458—462.
4. C. C. Alway, Crossing the desert, *Math. Gazette* 41(1957), 209.
5. O. Helmer, A problem in logistics. The jeep problem, *Project Rand Report* No. Ra15015, Dec. 1947.
6. R. Bellman, *Dynamic Programming*, Princeton Univercity Press, 1955, 103, ex. 54—55.
7. J. N. Fraklin, The range of a fleet of aircraft, *J. Soc. Indust. Appl. Math.* 8(1960)541—548.

[225]

[226]

① 原文为"red, white, and blue"．红、白、蓝是美国国旗所用的三种颜色．据说红色代表勇敢，白色代表真理，蓝色代表公正．看来这一传统源自欧洲，英国、法国、荷兰等国的国旗也是用这三种颜色．但现在说"红白蓝"一般即喻指美国国旗，从而在美国英语中有"国家"、"爱国"、"高贵"、"至高无上"等寓意．—— 译注

附录3 初等几何中的十九个问题

阿曼多·马查多(Armando Machado)

几年前，有人提出了初等几何中的一个问题．这个问题看起来十分幼稚，但是较为一般的方法却对它无能为力．它一定相当有名，因为它在数学圈里一再被提到．在图 A 3.1 中，我们有一个等腰三角形，其中 $\lambda=20^\circ$，$\alpha=60^\circ$，$\beta=50^\circ$，我们要算出 γ 和 δ 的值．

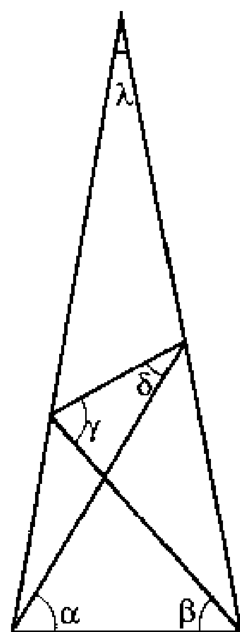


图 A 3.1

经过较为常规的计算后，我们得到了图中大多数未标出的角度，但是我们所需要的却一个也没有．我们得到的结论只是 $\gamma+\delta=\alpha+\beta=110°$．这时我感到虽然这个问题提得明确合理，但用初等几何的方法恐怕是不能解决的．于是我用一些三角学知识和一个袖珍计算器来确定出 γ 和 δ．令我吃惊的是，这些值十分整齐，具体地说，$\gamma=80°$，而 $\delta=30°$．对于这样的值应该存在一个初等的解决方法！我记得我采用了同样的 λ 和另一对 α 和 β（我记不起是哪一对了），结果对 γ 和 δ 我仍然得到整数值．我开始相信这是一个一般的现象，但是第三次试验告诉我情况并非如此．例如，对 $\lambda=20°$，$\alpha=20°$，$\beta=70°$，我们得到 $\gamma=2°.12201$，$\delta=87°.87799$．……我对 λ，α 和 β 试验了 $10°$ 的所有倍数，除去诸如 $\alpha=\beta$ 以及与前面的试例呈对称的那些平凡情况，能对 γ 和 δ 给出整数值的数据就是表 A 3.1 中列出的那些.（从现在开始，我将略去角度符号.）

[227]

λ	α	β	γ	δ
20	50	20	60	10
20	50	40	60	30
20	60	30	80	10
20	60	50	80	30
20	70	50	110	10
20	70	60	110	20

[228]

表 A 3.1

一个很自然的猜想是：在所有这些情况中，存在着一个解决这问题的初等方法，这个方法不涉及三角公式．确实，我的一位同事拉马尔霍（Margarita Ramalho）证实了这一点．有趣的是，对每一种情况，都不得不提出一个不同的证明，这是一个在数学中不很常见的现象．例如，对前两种情况的证明很平凡，虽然它们彼此不同．而第四种情况，也就是我们最初提出的那种情况，可

以用下述思路解决：把它的图形作关于垂直轴的镜像反射，再叠加在第一种情况的图形上，并注意这时出现的一个等边三角形和若干个等腰三角形.

最近，我正在赏玩数学软件 *Mathematica*，有人向我提出了同一问题. 我决定利用这个软件来寻找其他可能存在初等解法的初始数据. 我要求 *Mathematica* 对 λ 的每一个整数值以及 α 和 β 的每一个整数值或半整数值进行试验，从中选出那些相应的 γ 值和 δ 值为整数或半整数的情况. 这个实验给出了比以前多得多的可能有初等解法的候选者. 循着盖尔的一个想法，又有一些候选者露面了，其中包括了一些以 7 为分母的适当角度. 这些做法最终导致了表 A 3.2 的形成. [229]

问题	λ	α	β	γ	δ
1A	60/7	390/7	150/7	480/7	60/7
1B	60/7	390/7	330/7	480/7	240/7
2A	12	42	18	48	12
2B	12	42	30	48	24
3A	12	48	12	54	6
3B	12	48	42	54	36
4A	12	57	33	75	15
4B	12	57	42	75	24
5A	12	66	42	96	12
5B	12	66	54	96	24
6A	12	69	21	87	3
6B	12	69	66	87	48
7A	12	72	42	108	6
7B	12	72	66	108	30
8A	20	50	20	60	10
8B	20	50	40	60	30
9A	20	60	30	80	10
9B	20	60	50	80	30
10A	20	65	25	85	5
10B	20	65	60	85	40

(续表)

问题	λ	α	β	γ	δ
11A	20	70	50	110	10
11B	20	70	60	110	20
12	36	54	36	72	18
13A	45	45	15	52.5	7.5
13B	45	45	37.5	52.5	30
14A	360/7	240/7	120/7	270/7	90/7
14B	360/7	240/7	150/7	270/7	120/7
15A	360/7	345/7	150/7	435/7	60/7
15B	360/7	345/7	285/7	435/7	195/7
16A	72	39	21	48	12
16B	72	39	27	48	18
17A	72	42	24	54	12
17B	72	42	30	54	18
18A	72	48	24	66	6
18B	72	48	42	66	24
19A	72	51	39	81	9
19B	72	51	42	81	12
20A	120	24	12	30	6
20B	120	24	18	30	12

[230]

表 A 3.2

严格地说,并非所有这些情况都各不相同. 情况 3A,8A 和 12 有着相同的解法,而且它们是下述一系列情况的特例:λ 是满足 $0 < \lambda < 60$ 的任何一个值,$\alpha = 45 + \lambda/4$,$\beta = \lambda$,$\gamma = 45 + 3\lambda/4$,而 $\delta = \lambda/2$. 同样,情况 3B,8B 和 12 可以用相同的方法解决,而且它们是下述一系列情况的特例:λ 是满足 $0 < \lambda < 60$ 的任何一个值,$\alpha = 45 + \lambda/4$,$\beta = 45 - \lambda/4$,$\gamma = 45 + 3\lambda/4$,而 $\delta = 45 - 3\lambda/4$. 这样,表 A 3.2 列出的就是可能有着不同解法的 36 个问题. 不同问题的个数还可以进一步减少. 从表 A 3.2 中人们很容易发现一种对偶性. 对每个问题($\lambda,\alpha,\beta,\gamma,\delta$),我们

都可以联系上一个对偶问题$(\lambda',\alpha',\beta',\gamma',\delta')$，其中$\lambda'=\lambda,\alpha'=\alpha,\beta'=\alpha-\delta,\gamma'=\gamma,\delta'=\alpha-\beta$(问题12是自对偶的). 于是，一旦我们证明了这种对偶性，我们面临的就是18个问题，而对这种对偶性的证明，就是第十九个问题. 它们一起组成了本文标题中所说的19个问题. 关于对偶性的证明可以用如下思路解决：把对应于$(\lambda,\alpha,\beta,\gamma,\delta)$和$(\lambda',\alpha',\beta',\gamma',\delta')$的图形中的一个作关于垂直轴的镜像反射，再把它们叠合起来，然后应用射影几何中帕普斯定理的一个特殊情况. 我还没有用初等方法解决这些新情况中的任何一个(事实上，计算机给出的答案并不是真正的解，这一点在下面将看得很清楚). 在学术活动中的那些令人乏味的阶段——例如枯燥无聊的学会工作人员会议——考虑这些问题，或许是一种打发时间的好办法.

这也是人们在把计算机应用于数学时必须小心从事的一个有效例证. 我在使用软件 *Mathematica* 时，用的是其默认精确度，它大约给出6位准确数字. 而我的程序在检验一个数是否整数或半整数时所用的方法相当于考察其倍数的前5位小数. 除了表 A 3.2 中的情况外，计算机还建议了表 A 3.3 中的那些情况.

表 A 3.3 中有几件奇怪的事. 首先，有三个δ值差不多是整数或半整数，但它们的准确值并不是这样. 其次，有六个对偶问题遗失了. 这可能是一些舍入误差所造成的，于是我使用了 *Mathematica* 可在任一固定位数的精确度上工作的功能，以50位有效数字的精确度检验了表 A 3.2 和表 A 3.3 中的每一个值. [231] 表 A 3.2 中的每一个值仍保持正确，但表 A 3.3 中所有的值，以及它们那遗失的对偶问题，仅以近似的整数或半整数出现，虽然其近似程度非常好. 例如，表 A 3.3 中的第一个δ值变成了

6.50000 16063 95883 45352 05203 60429 84883 68773……

当时，我被这看来是不可思议的一个意外巧合弄得大吃一惊：几个在5位小数精确度下是半整数的计算结果竟然不是真正的半

λ	α	β	γ	δ
5	74.5	49	117	6.5
5	74.5	68	117	25.5
8	33.5	5.5	35	4
8	33.5	29.5	35	28
23	46	15.5	53	8.5
23	46	37.5	53	30.5
35	33.5	18	37	14.5
39	33.5	11.5	36.5	8.5
39	33.5	25	36.5	22
41	67	13	79.5	0.500003
59	33.5	4	35	2.5
59	33.5	31	35	29.5
61	32.5	7.5	35	5
61	32.5	27.5	35	25
67	32	4	33.5	2.5
67	32	29.5	33.5	28
68	46	28.5	62	12.5
68	46	33.5	62	17.5
68	55.5	41.5	95.5	1.5
77	49.5	14	62.5	0.999998
78	44	25.5	60.5	9
97	23	1.5	23.5	0.999998
97	23	22	23.5	21.5
129	21.5	7	26	2.5

表 A 3.3

整数. 其实,这里没有必要吃惊:计算机试验了大约 950000 个三[232]元组,这远远多于五位数组的个数.

附录4 真实性,只是真实性

数学的定律一旦要说到现实,它们就不是确定的;而它们要保持确定,它们就不能说到现实.

——爱因斯坦(Albert Einstein)

在本章中我写的是关于数学的一个特殊方面. 不过,在表述上有意地采用了非正式和非严格的方式. 爱挑剔的读者将发现有大量的漏洞好捉,但如果我到处都完全严格地来处理,就会妨碍思想的流畅性. 所以我希望读者对我有限地使用一下诗人的执照不要有所抱怨.

当我开始写这一章的时候,我的想法是,主要通过例子,对数学哲学的一个有争议的领域不带我个人偏向地给出一个粗略的阐述. 但在写作的过程中,我发现我至少在一些争议点上表明了自己的态度. 我决定在最后几段中给出我的结论. 我知道有人对这些事情思考了好多年,其深度很可能远远超过我,如果他们中有人不同意我的结论,我相信我或许其他读者都会很高兴读到他们的反驳文章.

数学结果的意义

数学以这样的特性在纯理论学科之林中独树一帜:只有它以绝对的确定性认定事实. 一旦一个定理被正确地证明了,那 [233] 就不可能有“另一方面”. 举个例子来说,每一个正整数都可以

写成至少四个平方数之和,那就是这么回事了. 在我看来,数学结果的这种无可争议性延伸到了数学文化的其他方面. 例如,关于这门学科中哪些结果是重要的,数学家们一般都有一致的意见,如代数基本定理、柯西(Cauchy)定理、中心极限定理、哥德尔不完全性定理(你或许可以建立你自己的有史以来十大重要定理表). 关于当代哪位数学家在做着最有意义和最激动人心的工作,数学界也有着一种公正的一致认识. 因此,在这一片和谐融洽的海洋之中发现在这门学科中的一个方面,数学家们是如此尖锐地发生冲突,真让人有点新鲜感. 这争论不是关于数学结果的正确性或深刻性,而是关于数学结果的**意义**,或更具体地说,是关于它们同现实的联系.

让我从一些实例开始. 我们知道不可能以经过那七座桥的每一座正好一次的方式游遍哥尼斯堡城. 而且我们获得这个知识的方式是通过推理,而这推理当然是数学的. 我们还知道不可能只用直尺和圆规来三等分一个任意角,而这里的推理过程要深刻和精妙得多. 我们知道在一个直角三角形中有 $a^2 + b^2 = c^2$(这在我的十大定理表上),但是直到最近我们还不知道当指数大于2时这个方程的类似物是否有整数解. 不过,我们中大多数人相信,这是出于我们的无知,而**在现实中**这样的整数解或者确实存在或者确实不存在.

另一方面,有些数学结果显然在现实中没有对应物. 比较一对著名的分解定理很能说明问题. 第一个是波尔约-格温(Bolyai-Gerwin)定理,它是说如果两个多边形 P 和 P' 具有相同的面积,那么我们就可以把 P 分割为有限多块(事实上是一些三角形),再把它们重新拼合起来而得到 P'. 这个定理当然有真实的对应物,比方说在拼图游戏的构造中. 我有一个计算机程序,可以让使用者在显示器屏幕上多少有点随意地画出一个三角形,机器随即把它切开,再把切成的块重新拼合,形成一个同样面积的正方形. 你不可能得到比这更真实的感觉了. 作为

对比,巴拿赫－塔尔斯基定理对3维空间中任何两个**不同体积**的球B和B'给出了与波尔约－格温定理相同的结论.我想大多数人会同意这个定理在真实世界中没有对应物(更不用说一个正在运行的计算机程序了).以像波尔约－格温定理这样的结果为一方,以像巴拿赫－塔尔斯基定理这样的结果为另一方,我在这篇随笔中希望做的,就是探索这两方面之间的"现实性缺口",因为正如我说过的,在这些事情上存在着分歧极大的观点.我从引述两个这样的观点开始.而主题,可能有人已经猜到了,就是集合论.

斯马尔扬(Raymond Smullyan)在向一群非专业工作者做关于连续统假设的演讲①.在解释了现在已知这个假设及它的否定都与集合论的标准公理相容之后,他说道,"现在,一些非常形式主义的数学家说这意味着连续统假设既不是真的也不是假的.然而,我不同意这观点,而且大多数——许多——数学家也不会同意.我们想知道连续统假设是**真**还是**假**"(字体变化是我做的).他的态度是,我们现有的工具,比方说策梅洛－弗兰克尔集合论,不足以让我们得知在阿列夫零与连续统基数之间是否存在基数,但是他认为无疑是要么存在要么不存在. [234]

麦克莱恩(Saunders MacLane)[2]的观点截然不同.他说道,"逻辑学家对策梅洛－弗兰克尔集合论的信念显然是基于一种柏拉图式的想像,即在真实世界的什么地方存在着某种集合的宇宙,对这个宇宙他们的公理将是适用的……我认为这个关于一个真实的集合宇宙的柏拉图式观念是绝对的胡言乱语."②

即使在这对连续统假设的争议中不带偏向(暂时),我也认

① 据公共广播系统(Public Broadcasting System)的NOVA节目,"数学的神秘之旅"(*Mathematical Mystrey Tour*).——原注

② 绝对的胡言乱语大概是胡言乱语的一种强化形式,就像绝对收敛是普通收敛的一种强化形式那样.——原注

为斯马尔扬有一点是错的．大多数——或者许多——数学家相信连续统假设非真即假的说法是不符合实际情况的．事实上，除了那些工作于集合论领域的数学家外，我还得去找出一位信奉这个观点的数学家．就像有人说的，我们知道我们可以同时有欧氏几何和非欧几何．现在我们看到有不同类型的集合论——而这又有什么不对？另一方面，我跟一些集合论家交谈过，其中有一些——虽然不是全部——取柏拉图式的观点．在我有限的样本中，有两个人认为这个假设很可能是不成立的，就同哥德尔所认为的那样．确实，在他生命行将结束的时候，哥德尔曾经谈到这样一个观点：连续统的基数很可能是阿列夫二，他甚至还写了一些东西，其中试图构造一个系统，这个系统将把他的这个观点作为一个结果．最近一段时间，有人告诉我，一些年轻的集合论家相信连续统假设是真的．有着相当数量的态度未定者，还有许多人，甚至包括集合论家，认为这个问题没有意义．

真，假或无意义

我们当然是在同著名的排中律打交道．为明晰起先，如果一个陈述满足排中律，就让我们说它是**确定的**，而如果它不满足，就说它是**不确定的**．一个我想大多数人会认为是确定陈述的例子是，"克娄巴特拉[①]的血型是A型．"虽然我们肯定将永远也不会知道这个陈述是真是假，但显然它不是真就是假，不是假就是真．另一方面，考虑这样一个陈述，"如果高斯从未诞生，也会有其他某个人在19世纪结束之前证出代数基本定理"．虽然很清楚这里说的是什么，但是问这个陈述是真还是假是不适宜的．这个陈述在相当程度上是一种说话的态度方式．我们是不是把所有这样的条件陈述都看作是不确定的？那么对"如果我

① 克娄巴特拉(Cleopatra，公元前69～前30)，古埃及托勒密王朝末代女王，即著名的"埃及艳后"．——译注

没有给你寄那封信,你就不会收到它"又如何说呢?

我的看法是,甚至在通常的语言中,一个特定的陈述是确定的还是不确定的,也不是总能弄清楚的. 作为最后的一个实例,请考虑"当地球上的生命全部消失之后,太阳还将继续存在至少100亿年". 对于这样一个东西我简直无所适从. 物理学是如此疯狂的一门学科,恐怕连时间的特定观念,即年,也是不确定的!

回到数学和数学家. 看来在排中律这个问题上,看法各异,分歧极大,形成了一个宽阔的对排中律的信念谱. 在这谱的极右端,是那些在所有的论证中都拒绝使用排中律的人;而在左端,我们有柏拉图主义者,例如前面提到的集合论家,他们认为这个规律极普遍地成立. 对于我们这些"走中间道路者"来说,像费马大定理那样的事情不是真就是假[①]. 对于右翼分子来说并非如此,他们感到它很可能是不确定的. 另一方面,主流人群却相信连续统假设是不确定的. 这是为什么呢?

[235]

让我们在一个特定的背景下,从一个特定的集合即自然数集合 **N** 出发,来审视这个问题. 主流人群认为这样的一个集合是存在的,它具有许多有趣的性质,其中之一就是费马大定理在 **N** 中或是真的或不是真的. 接下来我们进而讨论一个有时记为 $p(\mathbf{N})$ 有时记为 $2^{\mathbf{N}}$ 的集合,它就是由 **N** 的全体子集所组成的集合. 连续统假设断言,$p(\mathbf{N})$ 的每一个不可数子集都与 $p(\mathbf{N})$ 等势. 因此问题是:为什么这位"大街上的数学家"[②]会认为这个陈述是不确定的,而费马大定理倒是确定的? 我正巧也在大街上,无意中听到了下面的对话.

柏拉图:你为什么认为连续统假设既不真也不假?

大街上的数学家(以下简称**数学家**):噢,哥德尔证明它

① 这篇文章是在费马大定理被证明之前6年写的.——原注

② 作者用以代表"走中间道路者"或"主流人群"的虚拟人物.——译注

相容于集合论的其余部分，而科恩（P.J.Cohen）证明它独立于集合论的其余部分.

柏拉图：你说"集合论的其余部分"是指什么？

数学家：噢，那是叫做策梅洛－弗兰克尔集合论的某种东西. 它是一组公理，我想.

柏拉图：你知道这些公理是什么吗？

数学家：嗯，不，不太知道.

柏拉图：你仔细读过哥德尔和科恩的证明吗？

数学家：不，那不是我的领域. 我研究数论. 我正在试图证明费马大定理.

柏拉图：哦，数论！那么你研究自然数喽.

数学家：一直在研究.

柏拉图：那么你是不是有时要考虑自然数的序列？

数学家：经常考虑.

柏拉图：你是不是思考过自然数的**全体**序列？

数学家：偶尔. 我通常给它以积拓扑，并把它记为$\sum$.

柏拉图：我知道了. 那么我想这个$\sum$在你心目中是个定义十分明确的对象. 你有时候考察$\sum$的子集吗？

[236]

数学家：哦，是的.

柏拉图：那么是不是可以说你研究这些子集的时候感到十分自在？

数学家：当然. 它们有一些是开集，有一些是闭集，有些是可测集，而有些则不是. 十分有趣.

柏拉图: 那么当你研究这些不同类型的子集时,你是不是总在策梅洛-弗兰克尔公理的框架内工作?

数学家: 哦不,当然不.我已经说过,我甚至不能准确地知道这些公理是什么,而且无论如何我认为这种形式公理在我或者其他大多数职业数学家的工作中不起什么作用.

柏拉图: 瞧,现在你把我真的弄糊涂了.你告诉我你在研究$\sum$的一般子集时感到自在,然而当我问你一个直接关系到这类子集的基数的问题时,你回答这答案是不确定的,既不真也不假.而当我问你为什么会这样感觉时,你告诉我你听传说有两位著名的数学家已证明这个假设既不能从一组特定的公理得到证明也不能从这组公理得到否证——就是那组据你所说在你所做的工作中无论如何"不起什么作用"的公理.难道就没有可能这只不过是那些公理的一个短处,而不是客观事实的一种不确定性?

数学家: 我觉得有点头痛.

柏拉图: 我没有更多的问题了.

听了所有这些,而且我自己作为一个大街上的人,我也感到有点糊涂.排中,还是不排中,这是个问题.我开始思考那些右翼的非排中者.可以肯定,甚至他们也一定会同意某些陈述是确定的.例如下面这个陈述:

(1) π的第100^{100}位十进制小数是一个7.

好,现在存在一种算法可以解决这个问题(实际上有许多算法,而且其中有一些当前正在积极地运行着;见[3]).然而,看来我们肯定将永远不会得知对(1)的答案,因为得到它所需要的

时间很可能超过了预计的地球生命持续期(如果不是这样,就在那指数上再加几个零). 但是在原则上这计算是可以进行的,而且我相信没有人会提出这样一种理论,其中主张一个陈述是否确定要依赖于验证这一点所需的时间长度.

接下来我想到下面这个陈述:

(2) π 的十进制展开包含着一串连续的 100^{100} 个 7.

[237]

在这一点上右翼分子会声称这个陈述可能是不确定的,因为虽然当这个陈述是真的时候,存在一个有限的程序(虽然冗长)能够在原则上验证它,但如果这个陈述是假的,就没有这样的程序了. 把真假对换一下,这种情况就同费马大定理的情况一样了. 如果费马大定理是假的,那么一个有限的程序就能对此予以证实,具体地说,找出适当的整数 a,b,c 及 n,并执行规定的算术运算. 如果这定理是真的,就不存在这样的有限程序.

于是我又迈进一步,考虑下面的:

(3) π 的十进制展开包含无穷多个 7.

初看之下,这像是一个比(2)要容易对付的陈述(存在着非常强的"经验证据"提示这个陈述是真的;还是见[3]). 但是从另一个角度看,陈述(3)是不确定性的一位比(2)更好的代表. 在这种情况下,这个命题的真或假都不能用有限的计算验证,这一点稍稍想一下就会明白. (在经典的未解决问题中,类似的情况有孪生素数猜想.)

虽然我们几乎肯定地将永远不会知道(1)是不是真,但是可以想像,我们可能在某一天知道对(2)和(3)的答案. 事实上,存在着一个著名的猜想,说 π 是一个**正常数**,这意思是说在十进制表示下或在其他任何 b 进制表示下,它的每个数字符号都呈一致分布,而且更一般地有,每一种长度为 n 的数字符号序列都以渐近密度 b^{-n} 出现. 如果有人证明了这一点,那么陈述(2)和(3)就是显然的结果. 因此,令下一个陈述是:

(4) π 是一个正常数.

现在，我要作出推测：(4)将**永远不会**被证明或者被否证. 事实上，让我提议：或许**不存在**对(4)的证明或否证，其意思正同不存在对连续统假设的证明或否证，(对不起，重复一下)即它用标准的策梅洛－弗兰克尔集合论是不可证明的一样. 这可能让一些人立即认为是牵强附会。他们会说，证明一个关于连续统的如星云般渺茫的命题是不可证明的，是一件有理由的事，但是说到关于一个定义明确的序列中的子序列密度的具体问题时，人们应预期这种问题可以通过或许非常繁琐但常规的论证而得到解决. 作为回答，我要指出，一些比(4)"实在"得多的陈述**已知**用策梅洛－弗兰克尔公理是不可证明的. 或许最好的例子是希尔伯特第十问题研究的一个副产品. 一种特定的整系数多项式 P，我想是 9 个变元，具有这样的性质：对这个多项式在自然数中有一个根(即一个正整数 9 元组 $n=(n_1, n_2, \cdots, n_9)$，使得 $P(n)=0$)不存在(用策梅洛－弗兰克尔公理的)证明或否证. 这个问题在某种意义上至少比费马问题简单. 费马方程 $a^n+b^n-c^n=0$ 要比多项式复杂，因为它的变元出现在指数上.

因此现在我的问题是这样：假定有人真的证明了(4)不能用策梅洛－弗兰克尔公理证明或否证[①]，那么你会得出结论它既不真也不假吗？如果你的回答为是，那么让我就(3)问同样的问题. [238] 我想，我的朋友，那位大街上的数学家肯定认为 π 的展开要么包含有限多个 7，要么包含无穷多个 7. 如果有人证明了这个陈述是不可证明的，他会改变主意吗？当几天之后我在大街上遇到他的时候，我决定向他问这个问题. 我可以看出他已从前几天同柏拉图发生的冲突中恢复过来，重又感觉良好了. 我将

① 应当承认，这种可能是一种高度的假设. 证明不可判定性的已知方法看来不适合于像(3)和(4)这样的陈述，但人们可以想像有一些尚未得知的方法会给出这个结果. ——原注

上面的问题向他提出，他马上就予以回答。

"当然，在 π 的展开中，要么有有限多个 7，要么有无穷多个 7，不管有没有人证明了它是不可证明的. 这与连续统假设完全是两码事. 就像那个叫什么来着的人说的，上帝给了我们自然数. 所有其他的，像集合和基数，都是人的创造——在这种情况下，就是像康托尔(Cantor)和策梅洛这样的人的创造."

(当我提到 π 不是一个自然数时，他正确地指出这与所讨论的内容无关. 我们事实上在讨论的是对每一个**自然数 n** 产生一位数字的算法. 这个特定的算法来自对 π 作展开这件事显然是不相干的.) 我执意问他(3)不能用策梅洛－弗兰克尔公理证明这件事是不是在他心中产生什么问题. 他带着几分不耐烦答道，"公理啊，公理！我知道有许多事情不能用那些该死的公理证明. 柏拉图对此是正确的."

我还不打算让他脱身，因此我又问，"如果证明了 π 的正常性在策梅洛－弗兰克尔公理下是不可判定的，那会怎样？这里涉及到的陈述是关于子序列的渐近密度等概念的. 是上帝还给了我们那些东西，还是它们是人类的发明？"

他对我看了一会儿，然后说道，"我得回来再同你讨论这个问题." 遗憾的是，从那以后，我从未遇到过他.

无论怎么说，这里有更多的思想食粮. 人们应该在哪儿画界线？在哪一点上帝离去，人类接手？现实与胡言乱语之间的界线在哪儿？

现　　实

这时我自己也开始犯某种头部症状，于是决定让这件事停一段时间. 但是几天之后，我偶然读到了克努特(Donald Knuth)的文章《什么是一个随机序列？》(What is a random sequence?)[2]，于是这问题又到处跟随着我了. 我回过来再次思考陈述(4). 看来 π 不仅仅是一个正常数，它还满足比这强得多的条件. 例

如,如果我们考察这些小数位上的数字:$d_1, d_4, \cdots, d_{n^2}, \cdots$,它们似乎也呈一致分布. 对于序列($d_{n^3}$)无疑也有着同样的情况. 我们如今有了大量的数据(根据最新的计算,我们知道了20多亿位数字). 这提示我们应该强化定义,论及**超正常数**,它们不仅仅是正常的,而且它们的"可描述的"材质①也是正常的.(显然我们不能要求**所有的**子序列都是正常的,因为如果这样,给出任意一个数,我们可以挑选出一个子序列,它的各项就是出现7的数位,于是就没有什么超正常数了.)

幸运的是,对于**可描述的**一词,如今很可能大多数数学家和所有的计算机科学家都知道可以给出一个精确的定义. 正确的术语是**递归的**. 人们应该把这种可描述的序列看作是存在一种算法能算出它每一项的序列. 关键的事实是只存在可数多个这样的序列,因为只存在可数多种算法. 于是让我们称一个数是 [239] **超正常的**,当且仅当它的所有可描述子序列是正常的. 有一件事立刻就清楚了. π 不可能是超正常数,因为 π 的展开中出现7的数位显然形成了一个可描述的子序列(我刚刚描述了它),而相应的数0.777…是高度异常的. 在直觉上,我们把一个二进制的超正常数看作是人们通过重复抛掷一枚质量均匀的硬币而得到的序列. 如果在1600万次抛掷中,每当 π 的相应十进制小数位上是7时这硬币落下来就正面朝上,那一定十分奇怪,如果不奇怪倒奇怪了. 问题是超正常数**如果存在**,则一定是"不可描述的". 但是它们**确实**存在. 事实上,**非**超正常数的集合是一个零测度集! 而且,超正常数的集合是一个讨人喜欢的集合,准确地说,是一种 $\mathscr{F}_{\sigma\delta}$ 型集. 这个结果的证明本质上归功于博雷尔,它并不难,而且所涉及的知识也没有深过某种形式的大数律.

我发觉这是一个非常奇怪的结果,不是因为这里的数学,而

---

① 原文为substances,看来即"子序列". 如果不是笔误的话,那么这就是作者说过的他有限地使用一下诗人执照的例子.——译注

是因为它看来所意味的东西．如果我们想像重复抛掷这枚质量均匀的硬币，我们得到的序列将几乎肯定对应于一个超正常数，从而根据定义，它是一个不可能被描述的序列．这里的数学是清楚的，但是问题在于，用麦克莱恩的话来说，这样的序列是"存在于真实世界"还是它们只是我们数学想像力的臆造物？我不知道柏拉图对此会怎么说，但是我形成了我自己的一些结论．

我认为连续统既不是真的也不是假的，但不是由于哥德尔和科恩的结果，虽然这些结果大大增强了我的信念，而且没有它们我也不会产生这个信念．为解释我的态度，让我们回到自然数集合 **N** 和 **N** 的全体子集的集合 p(**N**)——重复一下，**N** 的**全体**子集——这个特定的背景．这个"全体"就是肇事者．在"现实"中，我们没有人曾经看到过一个无穷集合．我们看到过大量的有限集合，而且我们当然知道我们说这种集合的全体子集意思是指什么．大多数数学家也**认为**他们知道他们说**任何**集合——有限的或无穷的——的全体子集的集合意思是指什么．甚至我的朋友，那位大街上的数学家，带着他的Σ也在其中．然而且慢！当我们看到**全体**这个词同集合论联系上了的时候，我们心中应该亮起一盏红灯．康托尔认为考虑**全体**集合的集合是可以接受的，而我们知道这导致了什么样的困难．迄今为止，一个集合的全体子集的集合尚未导致任何矛盾，但是我的论点是，它标志着数学同现实失去联系的转折点．

左翼集合论家们，其中包括伟大的哥德尔，认为**在现实中**存在着一个特定的集合，它就是自然数集合（他们中有些人认为还有其他集合）的全体子集的集合．那么，或者是他们知道一些我不知道的东西，或者是他们在自己欺骗自己，把在有限情况下清楚的并可验证的某些东西外推到了无穷的情况，而在那里，依我的观点，这个概念很可能是没有意义的，正同问如果高斯从未诞生会什么情况那样没有意义．这并不是说 $p(S)$ 的概念在许多数学分支中不重要．我肯定我曾经用过它许多次．我说的只是

到这点上我们谈论的是数学而不是现实了. 因此对我来说,连续统假设是不确定的,其理由是这里涉及的集合是不确定的.

至于超正常数,我们还是必须从**全体**实数(或**全体**数字序列)的集合出发,把其中的一块扔掉,考虑余下的那块. 因为全体实数的集合令我神经紧张,我肯定不会把很多资本投到这个怪诞子集的存在上. 用另一种方式看,我们在谈论这样一种正面-反面序列:如果有人无穷多次地掷一枚质量均匀的硬币,我[240]们才几乎保证可以得到它. 换句话说,我们正在谈论如果某些不可能发生事情发生了才会发生的事情. 我发觉这甚至比谈论如果高斯未曾诞生将会发生什么情况**更不能接受**.

不幸的是,这不是这故事的终结. 我仍得面对 π 的展开中那无穷多个7的问题. 这一点归结为:当我们写下像"三点一四一五九点点点"这样的表达时我们是在做什么? 我觉得三,一,四,五,九和点让人觉得十分自在. 正是这"点,点,点"令我烦恼. 我想我知道我说这意思是指什么——但是我是这样吗?(我的阿斯匹灵药瓶在哪儿?)

那么所有这些将把我放在那关于排中律的信念谱的什么地方? 我会说在当中多少有点偏右的地方. 或许我应该称自己为头脑糊涂的保守党人. 无论怎么说,这是我现在觉得合适的方式. 当然,所有这些到下个星期可能会变化,特别是如果我正好在这期间遇上柏拉图.

参 考 文 献

1. D. E. Knuth, *The Art of Computer Programming Vol*. 2, Reading, Mass: Addison-Wesley, 1969.
2. S. MacLane, Are we all just specialists? *Mathematical Intelligencer* 8, 4(1986), 74—75.
3. S. Wagon, Is π normal? *Mathematical Intelligencer* 7, 3(1985), 65—67. [241]

关于本书

对那些被数学的抽象世界所深深吸引的人们来说，戴维·盖尔在《数学信使》上的专栏文章是快乐的一个基本来源. 这里，盖尔的专栏文章第一次结集成书.

这里的题材在数学领域中驰骋纵横，涉及广远，但是又经常回到深受喜爱的话题：三角形、铺砌、由简单递归关系给出的序列的神秘性质、游戏和悖论，以及一种特殊的自动机，它给出了这本集子的题目——《蚁迹寻踪》(*Tracking the Automatic Ant*).

书中讨论的问题有：

——为什么某些用分式定义的序列只产生整数?

——对一副有着无穷多张牌的纸牌连续进行简单的洗牌会把每张牌都带到最上面至少一次甚至无穷多次吗?

——沿着序列 1,2,4,8,…,128,256,…，我们得走多远才能保证其中一个数的开头数字是莎士比亚全部著作的一种编码?

——怎样才能让两人通过电话玩扑克，还要保证对手不作欺骗?

这些问题中，有一些有着简单却出人意料的答案，有一些仍然没有答案. 还有一些，虽然陈述起来很简单，但在一种严格的逻辑意义下，它们可能是“不可回答的”. 这本集子注重意外和神秘，分量大约相当.

这些小随笔层次各异，从像最经济的系鞋带方法这样的实际事情，到关于数学真理性质的半哲学思索.

不管你对数学的感觉如何,不管你是一位专业数学家还是一位普通读者,你会发现盖尔的作品既有教育意义又令人愉悦.

* * *

戴维·盖尔是加利福尼亚大学伯克利校区的一位数学荣誉退休教授. 从1991年到1996年,他作为《数学信使》的一名副主编,负责"数学娱乐"栏目. 他的专业领域包括对策论、几何学和组合学.

* * *

有关人士如是说:

"戴维·盖尔是一位富有创造性的数学家,他在《数学信使》上的专栏文章是快乐数学的一个丰富的源泉,这种快乐数学时常越界进入严肃数学. 把这些迷人的专栏文章结集成书让人阅读真是美妙极了. 这本书对正在日益扩增的趣味数学文献是一个重大的贡献."

——马丁·加德纳

"我每读一章都意外地发现一些新东西. 这大概是一件最完美的礼物,献给那些喜爱谜题、图案模式、证明、填装问题以及令人困惑的悖论的人们."

——约翰·J·康韦,普林斯顿大学

"我们向戴维·盖尔表示感激,他继承了马丁·加德纳的衣钵,出版了他在《数学信使》上那些生动活泼的专栏文章的集子. 现在它们将理所当然地进入更为广泛的读者群,并给他们带去快乐."

——理查德·盖伊,卡尔加里大学

"戴维·盖尔对数学中的奇闻异事可说饕口馋食,从历史经典到最新著述,从计算机生成的图案到你用

小棍在沙地上(或在你脑子里)的勾画，从举世万象到稀世珍宝，他百味皆纳．本书中各章内容精彩纷呈，令人流连忘返．其中大多数并未结束：因为那些深得个中精髓的读者将在放下此书后继续考虑回答盖尔的问题，并不断地构想出新的问题．”

——钱德勒·戴维斯，《数学信使》主编

译后记

我们生活在一个非数学的环境，可是盖尔说"到处是定理"．我们每个受过教育的人都学过数学，可是盖尔说"不存在这样的东西"．

我们平时经常看到物理的、化学的、天文的、地学的、生物的现象，但是对于数学，充其量只能看到涉及四则运算的"算术现象"．译者的工作部门可说是高学历人群的一个聚集地，我们中绝大多数人学过"高等数学"，但我们平时可以谈论一场足球赛、一部电视剧或一个热点新闻，而绝不会谈论数学．我想，即使是数学家，一旦离开了课堂、书斋或学术讨论会，也绝不会指望在平时听到有人谈论数学，他自己也不会在非正式场合讲述哈恩－巴拿赫定理、布劳威尔不动点定理，乃至——费马大定理！

这看来是很无奈的事，因为在某种意义上，数学是一种文化，一种精神产品，而且是一种在一般层次上难以表现、交流和传播的精神产品，因此它只能存在于创造和理解这种产品的人们的"心"中．盖尔说"到处是定理"，但它们好比是莎士比亚笔下的"树木间的谈话，溪流中的文章，石头中的喻示，以及每件事物中的益处"，需要我们用"心"去"发现"．盖尔说"不存在这样的东西"，这是因为数学在很大程度上是人类"心智"的"发明"．

当然，数学首先是一门科学，一门源于且用于人类实践活动的科学．而且，数学作为整个科学技术的基础，与科学技术的发展水平，从而与综合国力的发展水平，有着密切的联系．正是从

这个观点出发，我们需要普及数学．

然而，当我们说到数学的普及时，又回到了数学的文化方面．因为在我看来，数学普及，以至一般的科学普及，乃是一种用文化“包装”数学或科学使之让公众接受并在公众中传播的行为．当数学本身就具有丰富的文化内涵时，无疑就应该以揭示这种内涵为最好最恰当的普及手段了．综观世界上有影响的数学普及大师，如马丁·加德纳（Martin Gardner）、侯世达（Douglas R. Hofstadter）、斯图尔特（Ian Stewart）等，无一不是这样做的．

纵然如此，数学（通过其文化）的普及仍然是一件很困难的事．因为不同于其他的文化形式，数学文化具有深刻丰富的思想内涵却没有浅显直观的表现形态．换个角度说，对于普及者来说，即使在数学研究上颇有建树，但是最好还要有对数学思想在哲学层次上的领悟和对数学内容在传播方法上的技巧，才能把这种普及工作做得很理想；对于受普及者来说，则必须要具备一定的数学知识和一定的数学修养，才能很好地理解和欣赏这种文化．正是这个原因，数学普及在成效上一般来说落后于其他科学普及；也正是这个原因，存在于（相对）少数人“心”中的数学文化不能得到很好的表现、交流和传播，使我们总是处于一种非数学的环境．

不管怎么样，出于对数学事业的执著和挚爱，我们的数学家还是这方面做了不懈的努力，并且出了一批卓有成效的数学普及作家．除了上面提到的外，我们现在又有了戴维·盖尔．

盖尔的数学普及文章是有其特点的．

首先，盖尔在选择题材时除了注重通俗有趣外，还特别注重意外和神秘．所谓意外，即不符合常规思维模式的数学结果；所谓神秘，即目前难以预料结果的数学现象．当然，这两方面往往是联系在一起的，大多数题材是既意外又神秘．这个特点体现了盖尔对数学内容在传播方法上的技巧．

其次，盖尔在文章中经常把对具体数学事实的介绍引申到

自己对数学发展、数学思想方法、数学与现实的关系和数学本性的思索．从计算机对数学的影响，到数学证明中的变分方法和推广技巧，到数学在日常生活中的应用和对自然现象的解释（注意，这些应用和解释仍是在数学的文化层面上），直到对数学的中心概念不是数的阐释，可谓殚思极虑，沦肌浃髓．这个特点体现了盖尔对数学思想在哲学层次上的领悟．

这两个特点，使得盖尔的普及文章正如他自己所说的那样："读者将发现令人愉悦的东西，而且，也是富有启发性的东西．"

本书是盖尔在《数学信使》上的专栏文章的结集．严格地说，《数学信使》是一本数学圈子内的刊物，因此盖尔的专栏文章主要是针对数学家、大学数学专业师生以及与数学密切相关的科学教育工作者的．可以理解，这些文章中涉及的有些数学知识，不但超出了中学数学的范围，而且超出了大学理工科非数学专业和文科经济类专业的所谓"高等数学"、"工程数学"与"经济数学"的范围．

为了能让更多的读者阅读盖尔的文章，译者不揣冒昧，对凡是超出上述范围的数学知识，以译注的形式，或作粗略的介绍，或给出适当的参考读物．根据这个原则，比方说，译者对同余式运算作了注释，虽然这很初等，但超出了上述范围；而对罗尔（中值）定理却不作注释，虽然它比同余式运算要"高等"，但是在普通高等数学教程中必讲的．此外，对（古典）概率论和图论方面的知识一般也不作注释，因为前者是工程数学或经济数学中的必学内容，后者则十分直观明白．

但有一篇文章没有遵照上述原则给出数学知识方面的注释，那就是列于书末的附录4．这是一篇写给专业数学家看的文章，没有一定的数学修养是很难读通的．译注只能对数学知识略作介绍而不能对数学修养有所促进．如果对这篇文章中涉及的数学知识知之不多的话，恕我直言，还是以后再详读吧．不过，译者认为这是一篇很实在的文章．因此建议读者不管现在

还是将来,都要把它读一遍,一定受益匪浅.

译者还对一些历史文化和语言翻译方面的知识以及估计人们不太熟悉的数学家和数学定理作了注释.余下的注释,便相当于译者阅读本书的心得体会了.译者不怕见笑于同好,大胆说出自己的看法,也算是遵循华罗庚先生的教导:就是要“班门弄斧”.当然,目的是:若有讹谬,还望指教.

还有一点小小的遗憾,书中提到三位可能是华人(或华裔)的数学家,即 X.P.Kong,Fei Wang 和 Raymond Chen,按惯例应给出中文姓名写法,然而遍查未果.虽曾查到 Fei Wang 的 Email 地址(有好几个),并知道这位先生原毕业于天津南开高级中学,但发去的 Email 都被退回,原因是“接收者不详”.因此只好让他们的英文姓名原封不动.还有一位很可能是韩国人(或朝鲜人)的数学家 Yonghoan Kim,按韩国人取名习惯,有两种可能,一是“金龙焕”,一是“金容焕”.实无从查对,只好先取“金龙焕”.以上四位的中文(汉字)姓名,若有知晓者,还望告我,以在本书重印时补正.

在美国的日本数学家胜浦英文收到我求问他姓名汉字写法的 Email 时,正好在外,但连忙回复,说回到家中即用 Photoshop 画一个给我.几小时后我便收到了他的 tif 文件.在此表示感谢.

承上海教育出版社叶中豪先生委以翻译此书的重任,并费尽心力,匡我不逮,特表衷心感谢.此书翻译历时两年半有余.如此拖延,一是本职工作繁忙,业余时间较少;二是学识肤浅,时时受梗,苦思良久,方得其解.但是现在又发现一个原因:原来在盖尔的数学世界中左顾右盼,流连忘返.要不,当译完最后一个字时,怎么会有一种无名的惆怅呢?

译　　者

于上海西区半半斋